Contents

Review

Mobile Phone and Wearable Sensor-Based mHealth Approaches for Psychiatric Disorders and Symptoms: Systematic Review (e9819)
Jussi Seppälä, Ilaria De Vita, Timo Jämsä, Jouko Miettunen, Matti Isohanni, Katya Rubinstein, Yoram Feldman, Eva Grasa, Iluminada Corripio, Jesus Berdun, Enrico D'Amico, M-RESIST Group, Maria Bulgheroni.................................................................2

Original Papers

Use of Mobile and Computer Devices to Support Recovery in People With Serious Mental Illness: Survey Study (e12255)
Valerie Noel, Stephanie Acquilano, Elizabeth Carpenter-Song, Robert Drake.................................................................16

Believing Is Seeing: A Proof-of-Concept Semiexperiential Study on Using Mobile Virtual Reality to Boost the Effects of Interpretation Bias Modification for Anxiety (e11517)
Boris Otkhmezuri, Marilisa Boffo, Panote Siriaraya, Maria Matsangidou, Reinout Wiers, Bundy Mackintosh, Chee Ang, Elske Salemink.............21

Identifying Behaviors Predicting Early Morning Emotions by Observing Permanent Supportive Housing Residents: An Ecological Momentary Assessment (e10186)
Rajesh Nandy, Karabi Nandy, Emily Hébert, Michael Businelle, Scott Walters.................................................................34
Mobile Phone and Wearable Sensor-Based mHealth Approaches for Psychiatric Disorders and Symptoms: Systematic Review

Jussi Seppälä, MD, PhD; Ilaria De Vita, MSc; Timo Jääskä, PhD; Jouko Miettunen, MD, PhD; Matti Isohanni, MD, PhD; Katya Rubinstein, PhD; Yoram Feldman, MA; Eva Grasa, University of Oulu, Oulu, Finland

Department of Mental and Substance Use Services, Eksote, Lappeenranta, Finland

Ab.Acus srl, Milano, Italy

Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Oulu, Finland

Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland

Department of Diagnostic Radiology, Oulu University Hospital, Oulu, Finland

The Gertner Institute for Epidemiology and Health Policy Research, Tel Aviv, Israel

Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

CIBERSAM, Madrid, Spain

Fundació TIC Salut Social, Barcelona, Spain

m-RESIST, Barcelona, Spain

*these authors contributed equally

Corresponding Author: Maria Bulgheroni, MSc Ab.Acus srl Via Francesco Caracciolo 77 Milano, Italy Phone: 39 89693979 Email: mariabulgheroni@ab-acus.com

Abstract

Background: Mobile Therapeutic Attention for Patients with Treatment-Resistant Schizophrenia (m-RESIST) is an EU Horizon 2020-funded project aimed at designing and validating an innovative therapeutic program for treatment-resistant schizophrenia. The program exploits information from mobile phones and wearable sensors for behavioral tracking to support intervention administration.

Objective: To systematically review original studies on sensor-based mHealth apps aimed at uncovering associations between sensor data and symptoms of psychiatric disorders in order to support the m-RESIST approach to assess effectiveness of behavioral monitoring in therapy.

Methods: A systematic review of the English-language literature, according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, was performed through Scopus, PubMed, Web of Science, and the Cochrane Central Register of Controlled Trials databases. Studies published between September 1, 2009, and September 30, 2018, were selected. Boolean search operators with an iterative combination of search terms were applied.

Results: Studies reporting quantitative information on data collected from mobile use and/or wearable sensors, and where that information was associated with clinical outcomes, were included. A total of 35 studies were identified; most of them investigated bipolar disorders, depression, depression symptoms, stress, and symptoms of stress, while only a few studies addressed persons with schizophrenia. The data from sensors were associated with symptoms of schizophrenia, bipolar disorders, and depression. Although the data from sensors demonstrated an association with the symptoms of schizophrenia, bipolar disorders, and depression, their usability in clinical settings to support therapeutic intervention is not yet fully assessed and needs to be scrutinized more thoroughly.
Introduction

mHealth (ie, mobile health) is the intersection of electronic health and mobile devices for medicine and public health administration [1]. Many studies have actively exploited mHealth to provide questionnaires and qualitative feedback to facilitate treatment accessibility and participant retention or to monitor symptoms and treatment progress in a qualitative way. This is widely done using ecological momentary assessment (EMA) performed through e-diaries recording participants’ behavior. EMA collects self-report data through a variety of change-sensitive questionnaires [2-6]. However, self-monitoring has not always been shown to be a valid measurement of behavior. For example, a systematic review pointed out that electronic self-monitoring of mood among depression sufferers appeared to be a valid measure of mood in contrast to self-monitoring of mood among mania sufferers [7].

The rapid growth of smart-sensor integration in mobile phones and wearable devices has opened the prospect of increasing access to evidence-based mental health care. Mobile devices allow the collection of quantitative behavioral and functional markers in a transparent and unobtrusive way, providing an estimation of physiological and mental state [8-11]. A mobile phone-based approach may be valuable in gathering long-term objective data, aside from self-ratings, to predict changes in clinical states and to investigate causal inferences about state changes in patients (eg, those with affective disorders) [12].

In this review, the term sensor-based data includes the quantitative information supplied by the mobile phone and its embedded sensors. Information may range from acceleration to temperature and from light to pressure, but also from number of exchanged short message service (SMS) text messages to number of incoming and outgoing calls. Indeed, the variety of personal data, easily acquirable in this way, offers a unique opportunity to describe the person in terms of his or her lifestyle and behavior at the physical, cognitive, and environmental level [13,14].

Even if the evidence of association between sensor-based data and psychiatric disorder status and/or severity of psychiatric symptoms is limited and scattered [15-17], it is expected that appropriate management of these data may initiate a new trend in health care provision characterized by tailored and timely interventions [18].

Substantial treatment improvements have been achieved for several psychiatric disorders in the past decades. Nevertheless, the functional recovery of patients with schizophrenia is still low [19]. Treatment-resistant schizophrenia (TRS), especially, has a wide impact on the humanistic burden, which concerns patients and caregivers and involves several dimensions, such as quality of life, treatment side effects, caregiver burden, social impairment, suicide, violence, and healthy lifestyle [20]. Moreover, TRS patients show poor adherence to treatment-as-usual (TAU) intervention programs, which, in turn, cannot ensure continuity of assistance, immediacy of attention, tailored treatment, and caregivers’ integration [21]. In this context, the Mobile Therapeutic Attention for Patients with Treatment-Resistant Schizophrenia (m-RESIST) project [22] addresses patients with TRS by allowing caregivers and professionals to utilize mobile technology as part of the care process. These interventions determine a personalized flow of information based on a “Need 4 Help” scale and the stratification of patients depending on their risk level. m-RESIST is composed of three main parts: (1) a mobile phone connected to a smartwatch for patients and caregivers; (2) a Web-based dashboard for follow-up and monitoring by clinicians; and (3) a back-end system for managing data, interventions, and interactions between users [23].

The aim of this paper is to systematically review original studies on sensor-based data collection, targeting correlations between objective measurements of personal data and symptoms of psychiatric disorders to support the m-RESIST clinical approach. The main goal is to assess the perspective of integrated sensor-based mHealth interventions to deliver highly personalized mental care, monitoring the individual and his or her own modification along the way.

Methods

Overview

This systematic review has been performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [24]. Accordingly, strict eligibility criteria were applied in order to identify journal articles and reviews addressing the collection of sensor-based data in mental health and to investigate the association between sensor-based data and mental state. For a detailed description, see the PRISMA checklist in Multimedia Appendix 1.

Eligibility Criteria

Eligibility criteria are listed in Textbox 1.
Textbox 1. Eligibility criteria of papers to be included in this review.

- Types of participants: papers that studied participants with mental disorder diagnoses or symptoms of mental disorders (e.g., depression, anxiety, sleep disorders, psychotic disorders, stress, and panic disorders) were included; papers that studied participants without mental disorder diagnoses, but that analyzed participants with mental disorder symptoms (e.g., depression, anxiety, sleep disorders, and stress) were also included.

- Types of methods: studies reporting transparent and unobtrusive monitoring using commercially available wearable sensors (e.g., wristbands, bracelets, smartwatches, and mobile phones) were included. Studies describing Internet-based interventions, interactive voice-response technologies, and self-reporting interventions based on questionnaires without a sensor-based mobile app component were excluded. Furthermore, studies using obtrusive monitoring devices (e.g., chest band and helmets) were also excluded.

- Types of outcomes: studies reporting results associating mental health status and sensor-based data were included. Papers providing a description of the mobile app, but no statistical outcomes, were excluded.

- Language and time frame: English-language full-text articles, reviews, and conference abstracts were included in the review. Considering the trend of technology evolution, papers published between January 1, 2009, and September 30, 2018, were included.

Information Sources, Search Strategy, and Study Selection

The search for papers was performed using the following electronic databases: Scopus, PubMed, Web of Science, and the Cochrane Central Register of Controlled Trials. The following combinations of search terms were used: (“mental health” OR “mental disorder” OR depression OR anxiety OR psychosis OR schizophrenia OR “treatment resistant schizophrenia” OR bipolar OR insomnia OR stress) AND (mobile OR smartphone) AND (monitor OR sensing OR sensor).

Results of the search were made available in Excel files and included the title, authors, source, date, and abstract for study selection. Duplicated studies were removed before starting the selection. An eligibility check was performed on the title, keywords, and abstract of each study. Full-text copies of all potentially relevant papers, or papers where there was insufficient information in the abstract to determine eligibility, were obtained.

Study selection, according to the eligibility criteria described in Textbox 1, was performed independently by two reviewers: one with a clinical background and one with technological background. There were no cases of disagreements between the two reviewers.

The extracted information consisted of the following: (1) sensors that were used; (2) computed parameters; (3) participants (i.e., number and state of health); and (4) relation to clinical outcomes.

Results

As summarized in Figure 1, a total of 345 unique records were found from PubMed, 1038 from Scopus, 1358 from Web of Science, and 385 from the Cochrane Central Register of Controlled Trials, for a total number of 3126 hits. In all, 522 duplicates among the four databases were identified and removed.

A total of 1967 additional records were excluded because they reported on other technologies and/or other scientific fields. Another 226 were excluded because they did not report on suitable wearable sensors or did not report on sensors at all. An additional 234 were excluded because they described mainly methodological issues (e.g., protocols of analyses, mobile phone-based monitoring, and treatment apps) without suitable testing of subjects. Another 110 were excluded because they addressed pathologies, symptoms, and disorders outside of the mental health domain.

Altogether, 67 full-text papers were read; of these, 16 were excluded because they did not relate sensor data to health status assessment [25-40], while another six were feasibility studies with no relation to health status assessment [41-46].

In all, 35 articles were included in this review; two of them were complete reviews. One complete review addressed the association between a collection of behavioral features from mobile phones and wearable sensors with depressive mood symptoms in patients with affective disorders [47]. The other complete review addressed the use of digital health technology in the wider domain of serious mental illness [48]. Association of depressive mood symptoms with social behavior assessed through phone usage, physical activity measured through accelerometer and gyroscope, location measured by GPS, and overall device usage was not consistent across all studies [47,48]. The other 33 original papers are summarized in Table 1 [49-81].
Figure 1. Flowchart of study selection process.

Records identified through database searching (N=3126):
PubMed (n=345); Scopus (n=1038);
Web of Science (n=1358);
Cochrane (n=385)

Records excluded (N=2537):
other technological or scientific fields (n=1967);
no sensors or not wearable sensors (n=226);
only methodological issues (n=234);
other pathologies/symptoms/disorders (n=110)

Records after duplicates removed (N=2604)

Full-text articles excluded (N=22):
not relating sensor data to health status (n=16);
feasibility studies (n=6)

Full-text articles assessed for eligibility (N=67)

Studies included in qualitative synthesis (N=35):
reviews (n=2);
original studies (n=33)
<table>
<thead>
<tr>
<th>Source</th>
<th>Sample description</th>
<th>Collected data</th>
<th>Related clinical measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ben-Zeev et al [49]</td>
<td>47 healthy subjects</td>
<td>GPS, accelerometer, gyroscope, microphone, and light sensor</td>
<td>PHQ-9&lt;sup&gt;a&lt;/sup&gt;, PSS&lt;sup&gt;b&lt;/sup&gt;, and revised UCLA&lt;sup&gt;c&lt;/sup&gt; loneliness scale</td>
<td>Speech duration, sleep duration, and geospatial activity relate to PHQ-9; kinesthetic activity relates to UCLA loneliness scale.</td>
</tr>
<tr>
<td>Osmani V et al [50]</td>
<td>9 subjects with bipolar disorders</td>
<td>Accelerometer and gyroscope</td>
<td>HAMD&lt;sup&gt;d&lt;/sup&gt; and YMRS&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Psychiatric assessment scores relate to physical activity level at specific time intervals of the day.</td>
</tr>
<tr>
<td>Chow P et al [51]</td>
<td>72 healthy subjects</td>
<td>GPS</td>
<td>SIAS&lt;sup&gt;f&lt;/sup&gt; and DASS-21&lt;sup&gt;g&lt;/sup&gt;</td>
<td>Social anxiety and depression relate to time spent at home in specific time intervals of the day.</td>
</tr>
<tr>
<td>Boukhechba et al [52]</td>
<td>54 healthy subjects</td>
<td>GPS, phone calls, and messages</td>
<td>SIAS</td>
<td>Social anxiety relates to limited social life and reduced mobility.</td>
</tr>
<tr>
<td>Staples et al [53]</td>
<td>17 subjects with schizophrenia</td>
<td>Accelerometer and gyroscope</td>
<td>PSQI&lt;sup&gt;h&lt;/sup&gt;</td>
<td>Moderate correlation between sleep estimate and PSQI.</td>
</tr>
<tr>
<td>Sano et al [54]</td>
<td>66 healthy subjects</td>
<td>Accelerometer, gyroscope, skin temperature, skin conductance, phone calls, messages, and screen on/off</td>
<td>PSQI, Big Five Inventory Personality Test, MEQ&lt;sup&gt;i&lt;/sup&gt;, PSS, and MCS for mental health&lt;sup&gt;j&lt;/sup&gt;</td>
<td>PSQI and stress relate to phone usage.</td>
</tr>
<tr>
<td>Sano et al [55]</td>
<td>18 healthy subjects</td>
<td>GPS, accelerometer, gyroscope, skin conductance, phone calls, messages, and screen on/off</td>
<td>PSS, PSQI and Big Five Inventory Personality Test</td>
<td>Stress relates to phone usage and physical activities at specific time intervals of the day.</td>
</tr>
<tr>
<td>Stutz et al [56]</td>
<td>15 healthy subjects</td>
<td>Accelerometer, gyroscope, light, app usage, and screen on/off</td>
<td>PSS</td>
<td>PSS relates mainly to phone usage.</td>
</tr>
<tr>
<td>Difrancesco et al [57]</td>
<td>7 subjects with schizophrenia</td>
<td>GPS</td>
<td>Birchwood’s Social Functioning Scale</td>
<td>Locations detected through GPS relate well to the activities identified in the social functioning scale.</td>
</tr>
<tr>
<td>Osmani V [58]</td>
<td>12 subjects with bipolar disorders</td>
<td>GPS, accelerometer, gyroscope, and microphone</td>
<td>Mental scale (not specifically defined)</td>
<td>Physical activity and voice features relate to the patient’s state.</td>
</tr>
<tr>
<td>Renn B et al [59]</td>
<td>600 subjects with depression</td>
<td>GPS</td>
<td>PHQ-2&lt;sup&gt;k&lt;/sup&gt;</td>
<td>Limited association between mobility and depressive symptoms rating.</td>
</tr>
<tr>
<td>Mehrotra et al [60]</td>
<td>25 healthy subjects</td>
<td>Phone notification management (eg, clicks, decision, and response time), phone calls, and app usage</td>
<td>PHQ-8&lt;sup&gt;l&lt;/sup&gt;</td>
<td>Moderate correlation between depression state and notification management as well as phone and app usage in a 14-day period; limited correlation on shorter periods of time.</td>
</tr>
<tr>
<td>Grunerbl et al [61]</td>
<td>10 subjects with bipolar disorders</td>
<td>GPS, accelerometer, gyroscope, microphone, and phone calls</td>
<td>HAMD and YMRS</td>
<td>Good relationship between sensor data and the patient’s state.</td>
</tr>
<tr>
<td>Saeb et al [62]</td>
<td>28 healthy subjects</td>
<td>GPS and phone usage</td>
<td>PHQ-9</td>
<td>Good relationship between phone usage (ie, calls and duration) and depression symptoms as well as GPS processed data and depression symptoms.</td>
</tr>
<tr>
<td>Guidi et al [63]</td>
<td>1 patient with bipolar disorder</td>
<td>Microphone</td>
<td>QID&lt;sup&gt;m&lt;/sup&gt; and YMRS</td>
<td>No clear relationship between voice features and clinical assessment.</td>
</tr>
<tr>
<td>Beiwinkel et al [64]</td>
<td>13 subjects with bipolar disorders</td>
<td>GPS, phone calls, and messages</td>
<td>HAMD and YMRS</td>
<td>Phone usage relates positively to depression state while activity relates negatively to manic symptoms.</td>
</tr>
<tr>
<td>Wahle et al [65]</td>
<td>126 healthy subjects</td>
<td>GPS, accelerometer, and phone usage</td>
<td>PHQ-9</td>
<td>Depression symptoms relate to mobile phone extracted features.</td>
</tr>
<tr>
<td>Shin et al [66]</td>
<td>61 patients with schizophrenia, DSM-IV&lt;sup&gt;n&lt;/sup&gt;</td>
<td>Fitbit (ie, activity tracker)</td>
<td>PANSS&lt;sup&gt;o&lt;/sup&gt;</td>
<td>Psychiatric symptoms relate to lower activity level.</td>
</tr>
</tbody>
</table>
## Results

### Related clinical measures

<table>
<thead>
<tr>
<th>Source</th>
<th>Sample description</th>
<th>Collected data</th>
<th>Related clinical measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmius et al [67]</td>
<td>29 subjects with bipolar disorders and 20 controls</td>
<td>GPS</td>
<td>QID</td>
<td>Location recordings relate to depressive episodes.</td>
</tr>
<tr>
<td>Abrantes et al [68]</td>
<td>20 subjects with alcohol use disorders</td>
<td>Fitbit (ie, activity tracker)</td>
<td>PHQ-9</td>
<td>Physical activity correlates with reduction in the level of depression and anxiety.</td>
</tr>
<tr>
<td>Saeb et al [69]</td>
<td>48 healthy subjects</td>
<td>GPS</td>
<td>PHQ-9</td>
<td>GPS correlates with depression differently on weekdays and weekends.</td>
</tr>
<tr>
<td>Place et al [70]</td>
<td>73 subjects with at least one symptom of depression</td>
<td>GPS, accelerometer, gyroscope, phone calls, messages, microphone, and screen on/off</td>
<td>Semi-structured clinical interview</td>
<td>Physical activity and phone usage relate to depression symptoms.</td>
</tr>
<tr>
<td>Saeb et al [71]</td>
<td>206 healthy subjects</td>
<td>GPS, accelerometer, gyroscope, phone calls, messages, microphone, and screen on/off</td>
<td>PHQ-9 and GAD-79</td>
<td>No consistent relationship between GPS-based semantic location and depression or anxiety.</td>
</tr>
<tr>
<td>Faurholt-Jepsen et al [72]</td>
<td>61 subjects with bipolar disorders</td>
<td>Phone calls and messages</td>
<td>HAMD and YMRS</td>
<td>Significant correlation between depressive and manic symptoms and phone usage.</td>
</tr>
<tr>
<td>Sabatelli et al [73]</td>
<td>7 subjects with bipolar disorders</td>
<td>Wi-Fi-based position</td>
<td>HAMD and YMRS</td>
<td>Weak negative correlation between staying in clinics and self-reported state.</td>
</tr>
<tr>
<td>Rabbi et al [74]</td>
<td>8 healthy subjects (elders)</td>
<td>Accelerometer, gyroscope, barometer, and microphone</td>
<td>Friendship Scale, SF-36, CES-D², and YPAS¹</td>
<td>No clear relationship between sensor data and administered assessment scales.</td>
</tr>
<tr>
<td>Doryab et al [75]</td>
<td>3 healthy subjects</td>
<td>GPS, accelerometer, gyroscope, microphone, and light sensor</td>
<td>CES-D</td>
<td>Correlation between depression scales and sensor data.</td>
</tr>
<tr>
<td>Farhan et al [76]</td>
<td>60 healthy subjects</td>
<td>GPS, accelerometer, gyroscope, microphone, phone lock and unlock, light sensor, and phone call duration</td>
<td>PHQ-9</td>
<td>Correlation between PHQ-9 scores and all the sensor data is pointed out.</td>
</tr>
<tr>
<td>Canzian et al [77]</td>
<td>28 healthy subjects</td>
<td>GPS</td>
<td>PHQ-8</td>
<td>Significant correlation between mobility patterns and depressive mood.</td>
</tr>
<tr>
<td>Zulueta et al [78]</td>
<td>16 subjects with bipolar disorders</td>
<td>Phone keyboard usage</td>
<td>HAMD and YMRS</td>
<td>Accelerometer activity while typing, number of exchanged messages, and typing errors correlate with depression and mania scores.</td>
</tr>
<tr>
<td>Sano et al [79]</td>
<td>201 healthy subjects</td>
<td>Skin conductance, skin temperature, accelerometer, ambient light, GPS, phone calls, messages, app usage, and phone lock and unlock</td>
<td>PSS and MCS⁰</td>
<td>Skin conductance relates to MCS, skin temperature, and phone usage timing and duration; GPS relates both to PSS and MCS.</td>
</tr>
<tr>
<td>Tron et al [80]</td>
<td>25 subjects with schizophrenia, DSM-IV</td>
<td>Accelerometer, light, temperature</td>
<td>PANSS</td>
<td>Physical activity relates to PANSS.</td>
</tr>
</tbody>
</table>

**Notes:**

1. CES-D: Center for Epidemiologic Studies Depression Scale
2. SF-36: Short Form 36 Health Survey
3. CES-D²: Center for Epidemiologic Studies-Depression Scale, version 2
4. YPAS: Youth Psychiatric Assessment Scale
5. HAMD: Hamilton Depression Rating Scale
6. YMRS: Young Mania Rating Scale
7. PSS: Perceived Stress Scale
8. MCS: Mental Component Summary
9. PANSS: Positive and Negative Syndrome Scale
10. GAD-7: Generalized Anxiety Disorder 7-item Scale
11. QIDGPS: Querying the Impact of Depression and Generalized Panic Scale
12. PHQ-9: Patient Health Questionnaire 9-item Scale
13. Android activity-recognition API
14. Wi-Fi-based position
15. GPS, accelerometer, gyroscope, microphone, and screen on/off
16. Phone lock and unlock, light sensor, and phone call duration
17. Skin conductance, skin temperature, ambient light, GPS, phone calls, messages, app usage, and phone lock and unlock
18. Skin conductance, skin temperature, and phone usage timing and duration
19. Accelerometer, light, temperature
Interbeat intervals negatively correlate with positive symptoms; movement negatively correlates with negative symptoms.

Only five studies addressing schizophrenia were included in this review [53,57,66,80,81]. None of them included patients with treatment-resistant schizophrenia. Early studies by Ben-Zeev et al [34,49] analyzed patients’ location, activity, and speech, but did not associate sensor data to the severity of symptoms. Difrancesco et al [57] implemented a time-based method and a density-based method to identify the geolocations visited by 5 schizophrenic patients, detecting patients’ out-of-home activities with moderate recall. Staples et al [53] investigated sleep estimation of 17 patients by comparing the Pittsburgh Sleep Quality Index (PSQI), EMAs, and accelerometer data, but did not address the severity of symptoms. Psychiatric symptoms evaluated by the Positive and Negative Syndrome Scale (PANSS) among those with schizophrenia were related to lower activity level [66,80,81], while interbeat intervals correlated negatively with positive symptoms [81].

Most of the other included studies referred to depression [59,70] or symptoms of depression and anxiety in healthy subjects [49,51,52,54-56,60,62,65,68,69,71,74-77,79]. In one study, a limited association was found between mobility and ratings of depressive symptoms [59], while physical activity and phone usage were related to depressive symptoms in another [70]. In healthy subjects with symptoms of depression and anxiety, several data such as speech, sleep duration, mobility, and phone usage were related to severity of symptoms [49,51,52,54-56,60,62,65,74-77], while GPS-based semantic location did not correlate with depression or anxiety [71].

**Discussion**

**Principal Findings**

The data from sensors were associated with symptoms of schizophrenia, bipolar disorder, and depression. This may have the potential to change the nature of identification, follow-up, and treatment of mental disorders. Early identification of behavioral markers of psychiatric disorders may allow health care providers to react early to patients’ needs and deliver personalized dynamic treatment.

This systematic review uncovered a broad investigation, but still limited use, of data coming from mobile phones and wearable sensors to support therapeutic intervention for psychiatric disorders or for psychiatric symptoms. This review showed a high variability in participant selection criteria, investigation protocols, and data processing techniques, which limits the generalizability of the identified associations between...
sensor-based data and clinical assessment. This was also seen in three recent studies in the area of passive sensing in the mental health domain and the wider health care domain [13,47,82]. The available studies in this review often had methodological limitations (eg, small sample size, variations in the number of observations or monitoring duration, lack of randomized control group, and heterogeneity of methods).

In addition, there were issues related to usability of sensors and acceptance by patients; risks (eg, they may increase psychotic experiences and fears), feasibility (eg, psychiatric patients may have cognitive and economic limitations), risk-benefit ratio, costs, and health economics were not widely investigated. Also, potential biases in measurements due to the individual usage of the devices were only marginally addressed in most of the selected papers; for example, practical mobile phone use modalities (eg, only at work or at home) or reliability of wearable sensors (eg, a tight or loose smartwatch bracelet).

On the other hand, current psychiatric evaluation is strongly limited by assessment through scales on the day of the visit with the clinician and not necessary during a crisis (eg, “bad day” or relapse situation); it does not appropriately reflect the subjective experience of the patient nor the impact of the treatment in real life. The benefits of sensor-based data information may also be useful among those with TRS, as they show poor adherence to TAU programs of intervention; TAU intervention programs cannot ensure continuity of assistance, immediacy of attention, tailored treatment, and caregiver integration [21].

The data collected from sensors is expected to strongly contribute to behavioral monitoring and mental status assessment over time on an individual basis in a transparent way. Within an intraperson investigation, the data may be used as a trigger to personalized interventions facilitating the implementation of remote psychiatric therapeutic programs. It is expected that the long-term analysis of sensor-based data, building on a personal baseline and assessing individual modifications, may play a key role in clinical applications [14]. To realize this, all aspects of mobile phone sensor technology should be thoroughly investigated. Studies using rigorous methodology are needed to investigate the beneficial as well as the harmful effects of extracting behavioral markers of psychiatric disorders or symptoms from sensor-based data.

m-RESIST Project Contribution

Building on the results of this review, m-RESIST set up a framework to create a clinical decision support system (CDSS) based on a mobile therapeutic intervention for schizophrenic patients. The CDSS is designed to provide the users with necessary information to support health-related and clinical decision-making. The system utilizes available data sources in order to assess the patient’s condition using decision algorithms and, as a result, classify the clinical condition in order to provide clinical and lifestyle recommendations. The CDSS starts with a training period of two weeks, during which sensor-based data are collected, without activation of further system actions, in order to assess the patient’s baseline. Once trained, the system monitors the changes against the baseline. The functionality of the CDSS is based on the workflows developed by expert clinicians, reflecting the process of interaction between the system and its users in order to establish novel health care pathways. The CDSS activation is triggered by an event (ie, change in the baseline value) that is interpreted in a context of additional information that exists regarding a specific patient (ie, records in the patient’s file and information regarding attendance of scheduled visits) and a series of predefined conditions and actions [83].

The features supplied by sensor data that are used to trigger the CDSS are as follows: app number and duration of incoming, outgoing, and missed calls; number of incoming and outgoing SMS text messages by mobile phone; amount of time spent at home and in other places, measured by GPS data; and amount of time sleeping measured by physiological heart rate [83].

Conclusions

The data from sensors are associated with symptoms of schizophrenia, bipolar disorder, and depression, but their usability in clinical practice needs to be scrutinized more thoroughly. m-RESIST aims to support intervention administration by sensor-based data in TRS. m-RESIST also plans to go a step further in remote therapy management of TRS by implementing a CDSS to correlate clinical information and sensor-based data. In m-RESIST, a mental status evaluation based on the most common perceptions and risk behaviors of patients with schizophrenia has been developed, together with the usual clinical scales. A pilot study has been carried out and its results are under analysis.

Acknowledgments

This work was supported by the Horizon 2020 Framework Programme of the European Union (grant number 643552) and was partly funded by Fonds Européen de Développement Économique et Régional (FEDER) and Centres de Recerca de Catalunya (CERCA) Programme, Generalitat de Catalunya. We are grateful to all members of the m-RESIST project, who are also collaborative authors of this review under the name of m-RESIST Group.


http://mental.jmir.org/2019/2/e9819/
Sebú, Annika Seppälä, Jussi Seppälä, Valentina Simonetti, Matthias Stevens, Anna Triantafillou, Zsolt Szabolcs Unoka, Judith Usall, Vincenzo Vella, and David Vermeir.

**Conflicts of Interest**
None declared.

**Multimedia Appendix 1**
PRISMA checklist.

[DOCX File, 15KB - mental_v6i2e9819_app1.docx ]

**References**


22. m-RESIST. URL: https://www.mresist.eu/ [accessed 2019-01-29] [WebCite Cache ID 75nCphM9G]


Abbreviations

API: application programming interface
CDSS: clinical decision support system
CERCA: Centres de Recerca de Catalunya
CES-D: Center for Epidemiologic Studies-Depression scale
DASS-21: Depression, Anxiety, and Stress Scale
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
FEDER: Fonds Européen de Développement Économique et Régional
GAD-7: General Anxiety Disorder questionnaire
EMA: ecological momentary assessment
HAMD: Hamilton Depression Rating Scale
MCS: Mental Component Summary
MCS for mental health: Short Form-12 Physical and Mental Health Composite Scale
MEQ: Horne-Ostberg Morningness-Eveningness Questionnaire
m-RESIST: Mobile Therapeutic Attention for Patients with Treatment-Resistant Schizophrenia

http://mental.jmir.org/2019/2/e9819/ JMIR Ment Health 2019 | vol. 6 | iss. 2 | e9819 | p.14 (page number not for citation purposes)
PANSS: Positive and Negative Syndrome Scale
PHQ-2: Patient Health Questionnaire-2
PHQ-8: Patient Health Questionnaire-8
PHQ-9: Patient Health Questionnaire-9
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSQI: Pittsburgh Sleep Quality Index
PSS: Perceived Stress Scale
QID: Quick Inventory of Depression
SF-36: Short Form-36 Health Survey
SIAS: Social Interaction Anxiety Scale
SMS: short message service
TAU: treatment-as-usual
TRS: treatment-resistant schizophrenia
UCLA: University of California, Los Angeles
YMRS: Young Mania Rating Scale
YPAS: Yale Physical Activity Survey

©Jussi Seppälä, Ilaria De Vita, Timo Jämsä, Jouko Miettunen, Matti Isohanni, Katya Rubinstein, Yoram Feldman, Eva Grasa, Iluminada Corripio, Jesus Berdun, Enrico D’Amico, M-RESIST Group, Maria Bulgheroni. Originally published in JMIR Mental Health (http://mental.jmir.org), 20.02.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Background: Mental health recovery refers to an individual’s experience of gaining a sense of personal control, striving towards one’s life goals, and meeting one’s needs. Although people with serious mental illness own and use electronic devices for general purposes, knowledge of their current use and interest in future use for supporting mental health recovery remains limited.

Objective: This study aimed to identify smartphone, tablet, and computer apps that mental health service recipients use and want to use to support their recovery.

Methods: In this pilot study, we surveyed a convenience sample of 63 mental health service recipients with serious mental illness. The survey assessed current use and interest in mobile and computer devices to support recovery.

Results: Listening to music (60%), accessing the internet (59%), calling (59%), and texting (54%) people were the top functions currently used by participants on their device to support their recovery. Participants expressed interest in learning how to use apps for anxiety/stress management (45%), mood management (45%), monitoring mental health symptoms (43%), cognitive behavioral therapy (40%), sleep (38%), and dialectical behavior therapy (38%) to support their recovery.

Conclusions: Mental health service recipients currently use general functions such as listening to music and calling friends to support recovery. Nevertheless, they reported interest in trying more specific illness-management apps.

JMIR Ment Health 2019;6(2):e12255 doi:10.2196/12255

KEYWORDS
app; mental health; mHealth; mobile phone; serious mental illness

Introduction

The number of smartphones owned by people with serious mental illness (81%) has been approaching the rate of ownership in the general population recently (91%) [1,2], and the use of these devices by people with serious mental illness does not seem to differ substantially from that of the general population [3]. A survey of 457 people with serious mental illness found that a large majority have access to the internet via computers (89%), smartphones (54%), and tablets (32%) [4], which affords communication, socialization, and the opportunity to obtain information. Qualitative research has suggested that people with serious mental illness are interested in, and some may already be using, electronic devices without clinical team involvement to support their recovery [5]. For example, some people with serious mental illness reported using the internet to understand the precautions and side effects of medications; Instagram, to follow people who post daily positive messages; and YouTube, to watch videos that provide guided progressive muscle relaxation [5].

The concept of recovery began among mental health service users, but mental health personnel, including researchers, have
adopted the term [6]. Although recovery is an idiosyncratic concept, qualitative research has identified common themes: personal control over illness management, striving towards one’s goals, meeting one’s needs, and having a sense of responsibility [7-9]. People use routines and activities, such as employment and education, and engage with social support systems to promote their recovery [10]. Currently available technology may further empower people to manage their own recovery.

People with serious mental illness have recognized that technology could become a larger part of their recovery in the next few years [4,11]. Researchers are exploring the integration of smartphone and computer-based apps into psychiatric care for medication management, symptom monitoring, and shared decision making [12-14]. However, not all integrations have been adopted successfully by clients and mental health practitioners. Our study examined current use and interest in future use of specific features and apps of electronic devices by people with serious mental illness to support their recovery.

**Methods**

**Participants**

This study included a convenience sample of people with serious mental illness who were receiving mental health services from one small and one large mental health agency in New Hampshire. Of the 68 people who completed the survey, we excluded five participants from data analysis: two participants had inconsistent responses and three participants did not own an electronic device. The final sample consisted of 63 participants (31 men and 32 women). Participants ranged in age from 19 to 75 years (mean 41.6; median 42; SD 13.3), and the majority were white (84%, n=53) and never married (63%, n=40). The highest level of education attained by participants was some college or technical school education (38%, n=24), followed by a high school diploma or equivalent (30%, n=19). Nearly two-thirds of the participants were unemployed and not attending school (60%, n=38).

**Measures**

We developed a survey (Multimedia Appendix 1) based on findings from interviews of mental health service recipients [5]. This survey first assessed electronic device ownership and frequency of use. For those who owned a computer, tablet, or smartphone, we asked questions differentiating between general everyday use of these electronic devices and specific use for supporting recovery. The questions addressed several topic areas: frequency of use for general purposes, frequency of use for supporting recovery, ease of use, use of technology within mental health care, interest in trying new technologies, and interest in agency-based technical support services. We know from the qualitative study, which informed the development of this survey, that mental health service recipients use general/nonhealth apps (eg, Instagram, Facebook, and online games) to help them with recovery.

The participants rated whether their clinician discussed technology for supporting recovery, on a scale from 1 (no, never) to 10 (yes, at every visit). Participants also rated how comfortable they felt seeking/searching for help in using their electronic devices, on a scale from 1 (very uncomfortable) to 10 (very comfortable).

Because recovery is a highly individual experience [6], we allowed people to use their own understanding of the recovery process rather than an explicit definition. We pretested the survey for clarity and understanding with two consumers and revised the survey based on their feedback.

**Procedure**

The Dartmouth College Committee for Protection of Human Subjects in Hanover, New Hampshire, approved this study, which followed the principles outlined in the Helsinki Declaration. Over 4 weeks, we recruited participants from a community mental health center and a dual diagnosis treatment program in New Hampshire. The community mental health center, located in a rural area, serves approximately 1500 adults with serious mental illness each year. The city-based dual diagnosis treatment program serves between 30 and 40 men with co-occurring serious mental illness and substance use. A researcher and a research assistant approached clients in the waiting room and common areas of these centers, explained the study, and asked whether they would be interested in participating. In a few instances, the case managers approached the clients. Only clients who could provide informed consent and were receiving services at one of the two sites were eligible. We provided eligible, interested clients a tablet or paper-based survey, which included a description of the study and a consent statement to complete the survey. Consent participants completed the survey within 5 minutes. The researcher and research assistant were available to answer any questions and help those who requested assistance in completing the survey.

**Data Analysis**

We identified the five most frequently and least frequently used functions/apps for general everyday purposes. We then identified the five most frequently used functions/apps for supporting recovery. We also identified the top five functions/apps that participants were most interested in using to support their recovery in the future. We then used the Fischer exact test to identify whether these frequencies differed between the two sites. We used the Spearman rank correlation to assess whether age was associated with (1) the extent to which clinicians discuss technology with their clients and (2) the client’s level of comfort with seeking help for using electronic devices.

**Results**

**Ownership and Usage**

Approximately 90% of the people approached agreed to participate, but we do not have the exact number. The participants owned Android phones (67%, n=42), laptops/computers (63%, n=40), tablets (38%, n=24), and iPhones (29%, n=18). The most frequently used device by participants was Android phones (63%, n=40).

**General Everyday Use**

Participants regularly (ie, few times a week or every day) used their devices to access the internet (84%, n=53), make calls
(79%, n=50), text (79%, n=50), keep track of time (67%, n=42), access social media (60%, n=38), and track the weather (60%, n=38).

Participants rarely used their devices for mental health symptom monitoring (14%, n=9), sleep (14%, n=9), meditation (14%, n=9), dialectical behavior therapy (13%, n=8), physical symptom monitoring (eg, blood pressure and insulin level; 13%, n=8), or cognitive behavioral therapy (10%, n=6).

### Use to Support Recovery

To specifically support their recovery, participants most commonly listened to music (60%, n=38), accessed the internet (59%, n=37), called (59%, n=37), texted (54%, n=34), and used the clock feature to track time (41%, n=26). Participants’ frequency of use of the apps for either general every day or recovery purposes did not significantly differ between the two sites.

### Interest in Incorporating Technology Into Mental Health Recovery

Participants averaged a score of 3.2 (SD 2.6) on a scale from 1 (no, never) to 10 (yes, at every visit) when describing how frequently their case manager or clinician discussed the ways technology can support their recovery. The frequency of discussing technology with case managers or clinicians did not vary with age ($r_s$=−.19, $P$=.22). Among the two-thirds of participants (67%, n=42) who indicated that they would “probably” or “definitely” try new apps or technology to support their recovery, the most popular areas of interest included anxiety (45%, n=19), mood management (45%, n=19), mental health symptom monitoring (43%, n=18), cognitive behavioral therapy (40%, n=17), sleep (38%, n=16), and dialectical behavior therapy (38%, n=16).

A total of 48% (n=30) of participants found it easy to use new technologies; 29% (n=18) reported that it was sometimes easy and sometimes difficult to use new technologies. Participants reported that they either searched online or solicited help from family or friends when they need help using their device (75%, n=47). They rated their level of comfort with these approaches as 7.6 (SD 3.0) on a scale from 1 (very uncomfortable) to 10 (very comfortable). The level of comfort in seeking support did not vary with age ($r_s$=−.17, $P$=.27). Further, 60% of the participants indicated that they would “definitely” or “probably” work with an agency staff member who could help them use their devices, if such a person were available.

### Discussion

#### General Findings

Nearly all participants had access to devices that could connect to the internet. Between 40% and 60% identified specific features/apps they were currently using to support their recovery, namely, listening to music, accessing the internet, calling, texting, and keeping track of time. Two-thirds of the participants indicated that they were interested in trying new technologies to support their recovery. Participants were most interested in learning how to use apps that addressed anxiety, mood management, mental health symptom monitoring, cognitive behavioral therapy, sleep, and dialectical behavior therapy. Participants were moderately comfortable searching the internet or asking family or friends when they needed assistance using their device but were open to using technical support services if they were made available at the mental health center.

### Supporting Recovery

The majority of participants routinely used nonmental health features/apps, specifically those built automatically into electronic devices (eg, internet browser, texting apps, calling apps, and time tracking apps) to support their recovery. For example, one participant used the alarm on his phone for medication reminders, while another used the internet browser to learn more about mental health diagnoses [5]. In the present study, a substantial number of participants were interested in using mental health apps. These apps are publicly available; therefore, the following question arises: Why were the participants not using these apps? First, participants with low income have budget constraints that limit the brands of devices they can own and data plans they can afford, which directly impacts access to electronic resources [5]. Second, clients may view these apps as clinical tools that require support from a clinician. Our study participants reported minimal discussion with their clinicians about using technology to support recovery. Third, clients may not know where to find specific apps or how to decide on which ones to use. Mental health centers have a clear opportunity to involve a staff member with expertise in the field of mental health apps, such as a technology specialist, who can inform both clients and clinicians of vetted tools that may help support recovery efforts [5,15]. Evidence suggests that low-level support from professionals and the involvement of peers in a technology-supporting role would be helpful [16]. The majority of participants in this study were open to using such types of resources.

Between 38% and 45% of participants endorsed interest in apps in six target areas related to mental health, indicating that more than half of the participants are not interested in these apps. Consistent with the principle of shared decision making, researchers and clinicians could begin by taking advantage of the apps people are already using. Based on the study findings, participants found the apps that connect them to others or asking family or friends when they needed assistance using their device but were open to using technical support services if they were made available at the mental health center.

#### Limitations

Our study used a convenience sample in New Hampshire that lacked ethnic/racial diversity. We did not collect information on participants’ diagnoses. Behaviors described here were based on self-report, and people’s self-reported attitudes may not predict their behaviors.
Conclusions
People with serious mental illness use common features of smartphones, personal computers, and tablets to support their recovery, independent of the care they receive from mental health clinics. Clinicians and researchers may consider including a discussion of the apps clients are already using to monitor how effectively these tools support recovery efforts over time. A large minority of participants expressed interest in mental health-specific apps. Because the combination of interest, support, and acceptance is a key driver of adoption, clinicians and researchers may find successful adoption of these apps by starting with these clients and their choices rather than with all clients and specific apps.

Acknowledgments
The authors received funding for the research reported in this article from two nonprofit organizations: the Natalia Mental Health Foundation and the West Family Foundation. The funders had no role in the research activities or reporting of results.

Authors’ Contributions
VAN designed the study, participated in data collection, summarized the data, and led the write-up of the manuscript. SCA was involved in the study design, data collection, interpretation of the findings, and writing of the manuscript. ECS was involved in the study design and contributed to the interpretation of findings and writing of the manuscript. RED contributed to the interpretation of findings and writing of the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Survey on technology use.

References


©Valerie A Noel, Stephanie C Acquilano, Elizabeth Carpenter-Song, Robert E Drake. Originally published in JMIR Mental Health (http://mental.jmir.org), 20.02.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Believing Is Seeing: A Proof-of-Concept Semiexperimental Study on Using Mobile Virtual Reality to Boost the Effects of Interpretation Bias Modification for Anxiety

Boris Otkhmezuri1*, MSc; Marilisa Boffo2*, PhD; Panote Siriaraya3, PhD; Maria Matsangidou1, PhD; Reinout W Wiers2, PhD; Bundy Mackintosh4, PhD; Chee Siang Ang1*, PhD; Elske Salemink2,5*, PhD

1Faculty of Sciences, School of Engineering and Digital Arts, University of Kent, Canterbury, United Kingdom
2Faculty of Social and Behavioural Sciences, Department of Psychology, University of Amsterdam, Amsterdam, Netherlands
3Computer Science and Engineering, Faculty of Computer Science and Engineering, Kyoto Sangyo University, Kyoto, Japan
4Faculty of Science and Health, Department of Psychology, University of Essex, Colchester, United Kingdom
5Faculty of Social and Behavioural Sciences, Department of Clinical Psychology, University of Utrecht, Utrecht, Netherlands
*these authors contributed equally

Corresponding Author:
Boris Otkhmezuri, MSc
Faculty of Sciences
School of Engineering and Digital Arts
University of Kent
Jennison Building
University of Kent
Canterbury, CT2 7NT
United Kingdom
Phone: 44 1227 823 823
Email: bo201@kent.ac.uk

Abstract

Background: Cognitive Bias Modification of Interpretations (CBM-I) is a computerized intervention designed to change negatively biased interpretations of ambiguous information, which underlie and reinforce anxiety. The repetitive and monotonous features of CBM-I can negatively impact training adherence and learning processes.

Objective: This proof-of-concept study aimed to examine whether performing a CBM-I training using mobile virtual reality technology (virtual reality Cognitive Bias Modification of Interpretations [VR-CBM-I]) improves training experience and effectiveness.

Methods: A total of 42 students high in trait anxiety completed 1 session of either VR-CBM-I or standard CBM-I training for performance anxiety. Participants’ feelings of immersion and presence, emotional reactivity to a stressor, and changes in interpretation bias and state anxiety, were assessed.

Results: The VR-CBM-I resulted in greater feelings of presence ($P<.001$) and immersion ($P<.001$, $\eta^2_p=.74$) in the training scenarios and outperformed the standard training in effects on state anxiety ($P<.001$, $\eta^2_p=.3$) and emotional reactivity to a stressor ($P=.03$, $\eta^2_p=.12$). Both training varieties successfully increased the endorsement of positive interpretations ($P<.001$, $d_{repeated\ measures}=0.79$) and decreased negative ones ($P<.001$, $d_{repeated\ measures}=0.72$). In addition, changes in the emotional outcomes were correlated with greater feelings of immersion and presence.

Conclusions: This study provided first evidence that (1) the putative working principles underlying CBM-I trainings can be translated into a virtual environment and (2) virtual reality holds promise as a tool to boost the effects of CMB-I training for highly anxious individuals while increasing users’ experience with the training application.

(JMIR Ment Health 2019;6(2):e11517) doi:10.2196/11517

http://mental.jmir.org/2019/2/e11517/
KEYWORDS
anxiety; emotional reactivity; interpretation bias; cognitive bias modification; virtual reality; head mounted display; immersion; presence

Introduction

Background

Negative biases in information processing have been found to be a vulnerability factor and to play a causal role in the development and exacerbation of emotional disorders, particularly anxiety [1-3]. Empirical evidence has shown a robust relationship between anxiety and negative interpretation bias (for a review, see the study by Hirsch et al [3]). Whereas nonanxious individuals tend to favor positive or benign interpretations of ambiguous stimuli and situations, anxious individuals favor more threatening interpretations (ie, negative interpretative bias) and tend to exaggerate possible negative events in the future [2,4-7]. As a result, individuals vulnerable to anxiety experience more frequent and intense emotional reactions to everyday stressors and overestimate the presence of real threats in the environment.

Experimental research established that interpretation bias can be manipulated (or retrained) using a scenario-based procedure labeled Cognitive Bias Modification of Interpretations (CBM-I) [4]. In this training paradigm, participants repeatedly read short text scenarios describing ambiguous situations relevant to their type of anxiety, each one ending with a word fragment. The task of the participant was to read the text and resolve the word fragment in a meaningful fashion, which, when completed correctly, can result in a positive, negative, or neutral ending. The following is an example of a training scenario:

You’ve finished writing the answer to the second question in your exam.
You take a small break, looking at what’s left.
You then realize that the questions left are more difficult than you had anticipated. Checking the watch, you decide you’ve planned your time wll [well].

A subsequent question relating to the interpretation (eg, Will you have time to complete the exam?) is then presented, and a training-congruent answer (Yes or No) is positively reinforced (Correct). Then, the next trial starts with a new scenario, and so on.

In the positive interpretation condition, the ambiguous word stem can only be completed in a benign, anxiety-irrelevant way, and participants are thus trained toward positive resolutions of the described ambiguous scenario. In the control condition, typically an equal amount of positive and negative interpretations is presented. Two meta-analyses examining the effectiveness of CBM-I as a training intervention across anxiety and depression provided evidence for both near-transfer (ie, effects on interpretation bias measured with a similar task) and far-transfer (ie, effects on emotional reactivity to stressors and/or anxiety symptoms) effects, showing small to medium effect sizes [8,9], depending on the outcome. Anxious participants who were trained to consistently make benign interpretations of ambiguous information were more likely to generalize these more benign interpretations to new ambiguous stimuli or situations. As a result, participants showed lower levels of emotional vulnerability to stress, trait and state anxiety, and symptoms of anxiety [10-16], although the results are not consistent across studies [9]. Another recent meta-analysis looking at the effects of different types of Cognitive Bias Modification (CBM) interventions concluded that CBM effects were overall small or clinically nonrelevant [17]. However, greater beneficial effects on both anxiety and depression were observed for CBM-I paradigms compared with other types of CBM trainings.

Factors have been identified that impact CBM-I effectiveness. It has been shown that CBM-I training effects are stronger when participants actively process the (corrective) information [4]. In addition, imagining the described scenarios enhanced the training effects [18]. Furthermore, CBM-I effects have been found particularly pronounced when the training involves repeated practice over multiple sessions, indicating a dose-response relationship between the number of training sessions and effectiveness [9]. However, CBM-I training tasks generally include a very basic and unattractive layout (ie, a few lines of text presented on a neutral background), which makes training sessions highly unattractive. Participants who have undergone the training have reported it to be repetitive, boring, and monotonous [19,20]. The risk is that participants get easily distracted and stop being engaged with the training, resulting in less active processing of the content of the scenarios and the crucial training contingencies and, as a result, less learning [20]. Therefore, it is paramount to optimize the functional and aesthetic features of CBM-I training tasks so as to strengthen their beneficial effects and improve training adherence.

This Study

In this study, we tested the deployment of mobile virtual reality (VR) technology to transpose a scenario-based CBM-I training in a three-dimensional (3D) virtual environment, where the events described in the scenarios may virtually take place and be experienced in first person in a realistic fashion. The last two decades have seen an exponential increase in the use of VR technology in mental health treatment and within clinical research, with the greatest bulk of research showing the added benefits and long-term effects of virtual exposure therapy for different anxiety disorders, phobias, and post-traumatic stress disorder [21-25]. More recently, VR has also been extended to the adjunct treatment of psychosis, delivering cognitive rehabilitation, social skills training interventions, and VR-assisted therapies [26,27]. Furthermore, the development of information technologies has allowed mobile phones to meet all the requirements necessary to support VR (eg, appropriate central processor unit [CPU] and graphics processor unity [GPU] computing power, gyroscope integration) and, at the same time, to be portable and affordable. Their wide distribution allows more people to have...
access to immersive VR technology. In addition, the current generation of smartphones has become more than just devices for talking; the broadly available smartphones are capable of supporting 3D graphics, images, sounds, and software.

It is important to note that VR-based interventions and, in general, electronic health and mobile health interventions generally refer to the implementation of therapeutic principles in a digital environment rather than designing an entirely novel intervention paradigm [28,29]. In doing so, a mobile VR-based CBM-I training would harness the potential of simulating complex real-life environments where individuals can fully immerse themselves and explore, while keeping the effective principles underlying the training paradigm as intact as possible.

In VR, users are no longer simply external observers of images or text on a computer screen but are active participants immersed in a computer-generated 3D virtual world. By introducing specific perceptual cues evoking real-life contexts where (anxiety-relevant) ambiguous situations normally occur, VR strongly relies on the activation of the emotional reactions of the same ambiguous situation experienced in the real world and potentially increases the activation of relevant threat-related cognitive schemas [1]. The emotional experience, in turn, is related to presence, another important concept in VR, which involves the perception of the virtual environment as being real [30], creating the user’s sense of being in the VR environment. As such, “VR can be described as an advanced imaginal system: an experiential form of imagery that is as effective as reality in inducing emotional responses” [31,32].

The latter feature of VR is of special interest for the optimization of CBM-I training interventions, as the use of imagery instructions in CBM-I trainings has been found to boost their effects [9]. The ability of VR to “physically” immerse users within the ambiguous scenarios and to provide the proprioceptive perception of being an active agent in the virtual world has the potential to activate relevant memory schemas and evoke the typical interpretational and emotional response. Given recent insights into the importance of (a strong) discrepancy between expectations and the actual situation [33], VR may activate the dysfunctional schema and thus enhance the discrepancy with the positive interpretation provided in the CBM-I training, boosting prediction-error learning. As such, VR has the potential to enhance the therapeutic mechanisms underlying the training intervention. In fact, the activation of (anxiety-relevant) ambiguity and the related individual’s habitual pattern of biased information processing are necessary ingredients to successfully retrain it toward a more benign resolution [4,20]. Furthermore, the interactive and immersive properties of virtual environments may lead to an improvement of motivation to engage with the training application and the overall training experience compared with other media (eg, desktop computers).

Despite VR technology being used profusely as part of exposure therapy for anxiety disorders, the use of this technological platform in other forms of psychological interventions such as CBM training has received far less attention. To the best of our knowledge, only 1 proof-of-concept study has explored the feasibility of VR-based CBM training for social anxiety targeting attentional bias for threatening stimuli [34]. Although the study did not include a control group and was not designed, nor powered, to test the effectiveness of the intervention, the VR-based attentional bias training was associated with higher scores in enjoyment, flow, presence, and motivation than the standard training, indicating good acceptance and feasibility of the VR training intervention.

The main goal of this study was to examine the feasibility of using mobile-based stereoscopic 3D VR technology in a CBM-I training paradigm (virtual reality Cognitive Bias Modification of Interpretations [VR-CBM-I]) for performance anxiety to improve the users’ experience with the training program (ie, feelings of immersion and presence) and to potentially enhance training effects on state anxiety, emotional reactivity, and interpretation bias, compared with the standard training paradigm (standard CBM-I). We hypothesized that, compared with participants receiving the standard CBM-I training, participants completing the VR-CBM-I training would show (1) higher self-reported rates of immersion and presence in the training scenarios, (2) a greater endorsement of positive interpretations and less negative interpretations after the training, (3) a greater reduction in state anxiety after the training, and (4) lower emotional reactivity to stressors.

**Methods**

**Participants**

Participants were recruited through convenience sampling from the undergraduate student population of the University of Kent. Candidate students were invited by email to participate in a study on the use of VR to reduce anxiety levels. A total of 67 interested students aged 18 years and above were screened online for moderate-to-high trait anxiety (a score greater than 40 on the A-Trait subscale of the State-Trait Anxiety Questionnaire, STAI [35,36],—a standardized clinical measure of trait and state anxiety) and, when meeting this criterion, further invited to schedule a lab session. A total of 42 students (23 females and 19 males) aged between 18 and 35 years (mean 21.60 [SD 2.96]) with a mean trait anxiety score of 51.0 (SD 8.7) took part in the study.

**Procedure**

Upon arrival to the experimental laboratory, participants were briefly explained the goal and procedure of the study. Participants were informed that the study was focused on how CBM-I training can help support people with anxiety and that we were interested in exploring how different technologies, including VR, can facilitate the training of interpretation bias. The participants did not know the specific hypotheses of the study, nor that they would have received a stressor task. Participants were explained that they would be assigned to 2 groups of equal size, how a general scenario-based CBM-I training task worked, and that afterward they would complete general measures of stress, immersion, and system usability. After giving their informed consent, they were then assigned to either the standard CBM-I (n=21) or VR-CBM-I (n=21) training condition in a counterbalanced fashion, stratified by gender.
The experiment started with a baseline assessment of participants’ state anxiety (STAI A-State subscale) and interpretation bias (Recognition Task), followed by the training session completed on either the computer (standard CBM-I) or a head-mounted display system (VR-CBM-I) according to the allocated condition. At termination of the training and after an optional small pause, participants completed the post-training assessment of state anxiety (STAI A-State subscale), interpretation bias (Recognition Task), and perceived immersion (Immersion Experience Questionnaire [IEQ]) [37] and presence (Slater-Usoh-Steed [SUS] questionnaire) [38] during the training. The post-training assessment phase ended with a stress induction manipulation where participants rated their mood before and after performing a stressful cognitive test, the Anagram Stress Task, which has been designed to appear as an easy task to resolve but, in fact, being very difficult and almost impossible to complete, to assess their emotional response to actual failure. Finally, participants were fully debriefed about the study and the stressor procedure and compensated with a £10 voucher. The study was approved by the Research Ethics and Advisory Group of the Department of Engineering and Digital Arts of the University of Kent (reference number: 0631516).

Training Intervention

Standard Cognitive Bias Modification of Interpretations

The standard CBM-I training ran on a desktop computer on E-Prime [39], with scenarios presented as plain text on a white background (see Panel A in Figure 1). Scenarios were presented in 4 blocks of 10 scenarios each, with an optional break at the end of each block. Each scenario consisted of 3 lines that were ambiguous in terms of valence. The final sentence contained a missing word. After disappearance of the scenario, the omitted word was presented as a word fragment and disambiguated the scenario in a benign, anxiety-irrelevant way. Participants were instructed to complete the word fragment as quickly and accurately as possible by pressing the spacebar and typing the first missing letter. When not knowing the answer or after 10 seconds of inactivity, the correct answer was shown on the screen. A comprehension question then appeared and participants had to reply yes or no by pressing the Y or N button on the keyboard. Response accuracy and interpretation-relevant feedback were presented to reinforce the positive interpretation.

The scenarios were 40 event descriptions involving experiencing problems or potential failures in examination/test situations, which have been previously used in the performance anxiety domain [40]. An example of a (positive) scenario would be the following:

Together with a friend, you are preparing for a physics test. It’s the fourth time you are discussing a topic and your friend knows more than you. You think this is a ......

co-n-idence [coincidence]

Does your friend understand physics better than you? (Correct response: No)

It was just a chance that the friend knew more than you.

Figure 1. Representation of the standard Cognitive Bias Modification of Interpretation (CBM-I) and virtual reality Cognitive Bias Modification of Interpretation (VR-CBM-I) trainings: (A) Standard CBM-I training; (B) VR-CBM-I training (participant’s point of view on the computer room virtual environment on the top right corner); (C) Examples of virtual environments: classroom on the left side, living room in the middle, and book shop on the right side.
Virtual Reality Cognitive Bias Modification of Interpretations

The VR-CBM-I training was designed to be displayed on a commercially available VR head-mounted display system (Samsung Gear VR and Samsung Galaxy S6 smartphone). Overall, 7 narrative virtual environments were created in Autodesk Maya version 2017 and Autodesk 3ds Max version 2017 software programs to represent the same 40 training scenarios used in the standard CBM-I, with each environment combining 2 to 7 scenarios (eg, exam hall, classroom, and computer room; see Panel C in Figure 1 and Multimedia Appendix 1). The stereoscopic 3D virtual environments were then textured and rendered in Unity version 5.6.0f3, where the text in the training scenarios was added to the user interface. Participants could freely interact and explore the environment through head movements. The stories were presented as a pop-up text box appearing in the VR environment as soon as the participant started exploring it in the same presentation format as in the standard CBM-I training task (see Panel B in Figure 1). A voice recognition function that uses the inbuilt Android Speech Recognition function of the Samsung smartphone was developed and added to the VR-CBM-I system to allow participants to complete the word fragment by saying out loud the completed word and to answer the subsequent comprehension question by saying yes or no. To make the voice recording process easier and more understandable for the participants, a sound was used to indicate the start and end of the voice recognition process. Participants were able to repeat their answer again in case an incorrect word was given or recognized by the voice recognition system and were allowed to skip the step if they did not know the answer. To skip the step, participants needed to press the home button on the Samsung Gear VR glasses and the comprehension question then appeared. This design choice was made because of the different type of response input compared with the standard training, where the correct answer is normally shown automatically on the computer screen after 10 seconds when participants do not know the answer. Similar to the standard training, correct and incorrect answers to both the word fragments and the comprehension questions were visually highlighted in green and red, followed by the interpretation-congruent feedback in a pop-up text box.

Outcome Measures

Immersion and Presence Experience

Participants’ subjective experience of being immersed in the training scenarios was assessed with the IEQ [37], which consists of 31 items scored on a 5-point Likert scale covering 5 aspects underlying the immersive experience with a digital environment: emotional (6 items; eg, “To what extent did you feel that the scenario was something you were experiencing, rather than something you were just doing?”) and cognitive (9 items; eg, “To what extent did you feel you were focused on the scenario?”) involvement, which refer to the feelings and the amount of focus experienced while interacting with the digital environment; real-world dissociation (7 items; eg, “To what extent did you feel as though you were separated from your real-world environment?”), which refers to the sense of detaching from the outside world and increasing awareness of the digital environment; challenge (4 items; eg, “To what extent did you find the training scenario easy?”), which is the experience of being challenged by the digital environment; and control (5 items; eg, “At any point did you find yourself become so involved that you were unaware you were even using controls?”), which is the extent to which the user feels in control while interacting with the training. The IEQ was originally designed for the serious games field and has shown acceptable psychometric properties [41]. To adapt it to the context of this study, all game-related instances in the items were replaced with “involvement with the training scenarios.”

The experience of presence within the training scenarios was assessed with the SUS [38], a 6-item questionnaire rated on a 7-point Likert scale evaluating (1) the sense of being there in the scenarios as compared with being in a place in the real world (eg, “Please rate your sense of being in the scenario, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place. I had a sense of “being there” in the scenario.”), (2) how much the scenarios became the dominant reality (eg, “To what extent were there times during the experience when the scenario was the reality for you? There were times during the experience when the scenario was the reality for me...”), and (3) the extent to which a participant remembered the scenarios as a place visited, rather than as a computer-generated text or image (eg, “When you think back about your experience, do you think of the scenario more as ‘images’ that you saw, or more as somewhere that you visited? The scenario seems to me to be more like...”). Originally designed in the VR field, the questionnaire has been tested in multiple empirical studies and has shown to correlate with behavioral measures of presence [38,42]. For the purpose of this study, all VR instances in the questionnaire were carefully replaced with scenarios.

Interpretation Bias

Positive and negative interpretations were assessed with the Recognition Task before and after the training, a validated computerized task that has shown great sensitivity in capturing CBM-I training effects across both subclinical and clinical samples [4,12-14].

The task is similar in structure to the scenario-based standard CBM-I training (only with an added title), yet both the solution of the word fragment and the comprehension question do not disambiguate the scenario, which remained ambiguous. The task presented 10 new, unique ambiguous scenarios related to performance anxiety at each assessment timepoint. An example of a test scenario is the following:

Facts and Logic
You are working through a set of examples in your exam and concentrating very hard to try and remember the facts and logic you studied earlier.

When it comes to recalling what you have learnt you feel you know how effectively the test measures your true .....
memory ability

Was your memory for facts and logic being tested in an exam?

After presenting the 10 scenarios in a random order, the titles of the scenarios with 4 interpretations were presented again one at a time in random order. Participants were asked to rate the 4 interpretations on a 1 ("very different") to 4 ("very similar") scale for how similar in meaning each was to the original one [12,14]. The sentences represented (1) a possible positive interpretation, (2) a possible negative interpretation, (3) a positive foil sentence, and (4) a negative foil sentence. The 4 corresponding sentences of the Facts-and-logic-scenario are presented here as follows:

Facts and logic
1. You think you did not do well in the test because you cannot apply your good memory ability.
2. You think you will do well in the test because good memory is not important for it.
3. You think you will not do well in the test revealing your poor memory ability.
4. You think you will do well in the test because of your good memory ability.

Emotional Outcomes

State anxiety was assessed with the A-State subscale of the STAI questionnaire Form Y [35], which is a standardized measure of subclinical and clinical trait (A-Trait subscale) and state (A-State subscale) anxiety with very robust psychometric properties [36], including 20 items rated on a 4-point Likert scale. Stress reactivity to failure was measured by assessing participants’ emotional responses to a cognitive stressor, the Anagram Stress Task [43]. Participants were presented with 13 anagrams of different levels of difficulty that had to be solved within 28 seconds by typing the correct word. A new anagram was presented after responding or when the 28 seconds were expired. Participants were told that the task was a test of their language skills, which were found to be a reliable predictor of success in many domains, and that students normally perform well in such a task. Although the test appeared relatively easy, it was in fact extremely difficult, so that all participants failed most items. Before and after the task, participants rated how anxious and how sad they felt on 2 visual analogue scales (VAS) ranging from 1 ("happy" or "relaxed") to 100 ("sad" or "anxious").

Results

Sample Descriptives

Table 1 shows baseline sample descriptives. Comparison between the groups revealed no significant baseline differences in age, gender, trait and state anxiety, previous experience with VR, or accuracy in the solution of both the word fragments and the comprehension questions in the pretraining Recognition Task.

Presence and Immersion

An independent-samples t test was carried out to examine whether participants completing the VR-CBM-I experienced more intense feelings of presence during training than participants completing the standard CBM-I training, as measured by the mean rating on the SUS items. Results showed that the VR-CBM-I group experienced significantly higher levels of presence (mean 4.97 [SD 0.90]) than the standard CBM-I group (mean 3.33 [SD 1.30]; t_{40}=4.75, P<.001, d=1.47).

To test whether the VR-CBM-I condition was associated with a more immersive experience than the standard CBM-I condition, a multivariate analysis of variance was carried out using the 5 IEQ subscales. A significant main effect of Group was observed (F_{5,36}=20.9, P<.001, η^2_p=0.74), indicating that the VR-CBM-I group experienced a greater degree of immersion in the training scenarios than the standard CBM-I group. Univariate analyses indicated that the VR-CBM-I and standard CBM-I groups differed significantly on the following 4 subscales, control, real-world dislocation, emotional involvement, and cognitive involvement, and not on the challenge subscale (see Table 2).
Table 1. Sample descriptives at baseline: group means (SD) or frequencies (%), statistics, P value, and measure of effect size (Cohen $d$ or Cramer $V$).

<table>
<thead>
<tr>
<th>Variables</th>
<th>VR-CBM-I$^a$</th>
<th>Standard CBM-I$^b$</th>
<th>Statistics</th>
<th>$P$ value</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>7 (16.7)</td>
<td>12 (28.6)</td>
<td>0.24 (1)</td>
<td>N/A$^c$</td>
<td>.12</td>
</tr>
<tr>
<td>Females</td>
<td>14 (33.3)</td>
<td>9 (21.4)</td>
<td>0.24 (1)</td>
<td>N/A</td>
<td>.12</td>
</tr>
<tr>
<td>Age, mean (SD)</td>
<td>21.05 (1.91)</td>
<td>22.14 (3.7)</td>
<td>N/A</td>
<td>−1.20 (40)</td>
<td>.24</td>
</tr>
<tr>
<td>Trait anxiety, mean (SD)</td>
<td>50.43 (8.63)</td>
<td>51.57 (8.93)</td>
<td>N/A</td>
<td>−0.42 (40)</td>
<td>.68</td>
</tr>
<tr>
<td>State anxiety, mean (SD)</td>
<td>45.76 (6.30)</td>
<td>44.48 (4.66)</td>
<td>N/A</td>
<td>0.75 (40)</td>
<td>.46</td>
</tr>
<tr>
<td>Baseline accuracy recognition task, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word fragments</td>
<td>0.76 (0.77)</td>
<td>1.10 (0.77)</td>
<td>N/A</td>
<td>−1.41 (40)</td>
<td>.17</td>
</tr>
<tr>
<td>Comprehension questions</td>
<td>1.86 (1.46)</td>
<td>1.96 (1.28)</td>
<td>N/A</td>
<td>−0.23 (40)</td>
<td>.82</td>
</tr>
<tr>
<td>Previous experience with VR$^d$, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>4 (9.5)</td>
<td>3 (7.1)</td>
<td>0.17 (1)</td>
<td>N/A</td>
<td>.68</td>
</tr>
<tr>
<td>No</td>
<td>17 (40.5)</td>
<td>18 (42.9)</td>
<td>0.17 (1)</td>
<td>N/A</td>
<td>.68</td>
</tr>
</tbody>
</table>

$^a$VR-CBM-I: virtual reality Cognitive Bias Modification of Interpretations.
$^b$CBM-I: standard Cognitive Bias Modification of Interpretations.
$^c$N/A: not applicable.
$^d$VR: virtual reality.

Table 2. Mean scores for the Immersive Experience Questionnaire subscales (SD in parentheses), $F$ statistics, $P$ value, and effect size ($\eta_p^2$) for the VR-CBM-I and standard CBM-I groups.

<table>
<thead>
<tr>
<th>IEQ$^a$ subscale</th>
<th>VR-CBM-I$^b$, mean (SD)</th>
<th>Standard CBM-I, mean (SD)</th>
<th>$F$ statistics ($F_{1,40}$)</th>
<th>$P$ value</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenge</td>
<td>4.18 (0.79)</td>
<td>3.96 (0.80)</td>
<td>0.77</td>
<td>.39</td>
<td>0.02</td>
</tr>
<tr>
<td>Control</td>
<td>4.85 (0.86)</td>
<td>3.32 (0.84)</td>
<td>33.73</td>
<td>&lt;.001</td>
<td>0.46</td>
</tr>
<tr>
<td>Real world dislocation</td>
<td>5.14 (0.52)</td>
<td>3.03 (0.81)</td>
<td>100.33</td>
<td>&lt;.001</td>
<td>0.72</td>
</tr>
<tr>
<td>Emotional involvement</td>
<td>4.49 (0.95)</td>
<td>3.10 (0.77)</td>
<td>26.86</td>
<td>&lt;.001</td>
<td>0.40</td>
</tr>
<tr>
<td>Cognitive involvement</td>
<td>5.34 (0.57)</td>
<td>4.26 (0.70)</td>
<td>30.26</td>
<td>&lt;.001</td>
<td>0.43</td>
</tr>
</tbody>
</table>

$^a$IEQ: Immersion Experience Questionnaire.
$^b$VR-CBM-I: virtual reality Cognitive Bias Modification of Interpretations.
$^c$CBM-I: standard Cognitive Bias Modification of Interpretations.

Interpretation Bias

To test whether the VR-CBM-I training was more effective in changing interpretations than the standard CBM-I training, the Recognition Task data were subjected to a 2×2×2×2 mixed analysis of variance (ANOVA) with Group (VR-CBM-I vs standard CBM-I) as between-subjects factor and Time (pre- vs post-training), Valence (positive vs negative), and Interpretation type (Target vs Foil) as within-subject factors. A significant main effect of Interpretation type was revealed ($F_{1,40}=71.0$, $P<.001$, $\eta_p^2=0.64$), as well as 2 significant 2-way interaction effects (Time×Valence, $F_{1,40}=36.3$, $P<.001$, $\eta_p^2=0.48$; and Time×Interpretation type, $F_{1,40}=7.3$, $P=.01$, $\eta_p^2=0.15$). These effects were subsumed within a significant higher order 3-way interaction effect of Time×Valence×Interpretation type ($F_{1,40}=8.2$, $P=.007$, $\eta_p^2=0.17$). To decompose the 3-way interaction effect, separate analyses were carried out for target and foil sentences (Interpretation type) separately. Both analyses revealed significant Time×Valence interaction effects (Targets: $F_{1,40}=36.0$, $P<.001$, $\eta_p^2=0.47$; Foils: $F_{1,40}=10.0$, $P<.001$, $\eta_p^2=0.20$). The effects sizes for these interaction effects were larger for the targets compared with the foils, suggesting stronger training effects on interpretations than on foil statements. Subsequently, separate pairwise $t$ tests were conducted to decompose the Time×Valence effects for targets and foils separately. Consistent with the goal of the positive interpretation training conditions, there was a significant increase in positive target interpretations ($t_{41}=-5.1; P<.001; d_{rm}=0.79$; pretraining: mean 2.16 [SD 0.40]; post-training: mean 2.50 [SD 0.44]) and a significant decrease in negative target interpretations ($t_{41}=4.7; P<.001; d_{rm}=0.72$; pretraining: mean 2.44 [SD 0.44]; post-training: mean 2.07 [SD 0.50]). The effects
were less pronounced for the foils, and only the increase in the endorsement of positive foil sentences was significant ($t_{41} = −5.2, P < .001$; pretraining: mean 1.95 [SD 0.40]; post-training: mean 2.24 [SD 0.44]; negative foil sentences, $t_{41} = 0.3, P = .76$; d$_{m}$ = 0.04; pretraining: mean 1.95 [SD 0.48]; post-training: mean 1.93 [SD 0.48]). Collectively, this suggests that the stronger training effects on targets versus foil sentences are driven by the specificity effects in the negative interpretations. The 4-way interaction effect of Group x Time x Valence x Interpretation type was not significant ($F_{1,40} = 0.9, P = .35$, $η^2_p = 0.02$), indicating that the VR-CBM-I training did not result in stronger effects on interpretations than the standard CBM-I training.

**State Anxiety**

To test whether the VR-CBM-I training resulted in a stronger reduction in state anxiety than the standard CBM-I training, the STAI A-State scores were subjected to a 2 (Group: VR-CBM-I vs standard CBM-I training) x 2 (Time: pre- vs post-training assessment) mixed ANOVA. There was a significant main effect of Time ($F_{1,40} = 120.9, P < .001$, $η^2_p = 0.75$) and a significant Group x Time interaction effect ($F_{1,40} = 22.0, P < .001$, $η^2_p = 0.35$), confirming the stronger effects of the VR-CBM-I on anxiety. That is, although state anxiety did not differ significantly between the 2 groups before training ($t_{41} = 0.8, P = .46, d = 0.23$), participants who completed the VR-CBM-I training reported significantly less anxiety symptoms after training than participants in the standard CBM-I group ($t_{41} = −3.1, P = .003, d = 0.97$; see Panel A in Figure 2).

**Stress Reactivity**

To test whether the VR-CBM-I training resulted in a reduced emotional response to the stressor, the VAS Anxiety was subjected to a 2 (Group: VR-CBM-I vs standard CBM-I training) x 2 (Time: pre- vs post-stressor) mixed ANOVA. In addition to significant main effects of Time ($F_{1,40} = 12.9, P = .001$, $η^2_p = 0.24$; increase in anxiety from pre- to poststressor) and Group ($F_{1,40} = 15.4, P < .001$, $η^2_p = 0.28$; lower anxiety in the VR-CBM-I group), the predicted Group x Time interaction effect was significant ($F_{1,40} = 5.2, P = .03$, $η^2_p = 0.12$). Consistent with our predictions, the stress task resulted in a significant increase in anxiety in the standard CBM-I group ($t_{20} = −3.3, P = .003, d = 0.72$), but this was not the case for the participants who followed the VR-CBM-I training ($t_{20} = −1.4, P = .18, d = 0.31$; Panel B in Figure 2).

Exploratory, we examined whether the effects of training on emotional reactivity generalized to depressed feelings by subjecting the VAS Sadness to the same 2 x 2 mixed ANOVA. Again, significant main effects of Time ($F_{1,40} = 41.8, P < .001$, $η^2_p = 0.51$; significant increase in sadness from pre- to post-stressor) and Group ($F_{1,40} = 12.2, P = .001$, $η^2_p = 0.23$; lower sadness scores in the VR-CBM-I group) were observed. However, the Group x Time interaction effect was not significant ($F_{1,40} = 2.7, P = .09$, $η^2_p = 0.07$).

**Posthoc Analyses**

To examine whether the observed changes in state anxiety and emotional reactivity to the stressor were associated with perceived immersion and presence, Pearson correlations were computed between the IEQ and SUS scores and changes in state anxiety over the course of the training and changes in anxiety reactivity due to the stressor (see Table 3). Change indices were calculated by subtracting pretraining from post-training scores (ie, negative values indicate greater decrease). Stronger reduction in state anxiety across the training was significantly correlated with higher control, real-world dislocation, emotional involvement, and cognitive involvement. Furthermore, less anxiety reactivity was significantly correlated with greater perceptions of real-world dislocation and cognitive involvement.

**Figure 2.** (A) Mean (and SE) state anxiety scores from pre to post-training for the virtual reality Cognitive Bias Modification of Interpretation (VR-CBM-I) and standard Cognitive Bias Modification of Interpretation (CBM-I) groups. (B) Mean (and SE) VAS anxiety scores from pre- to post-stressor for the VR-CBM-I condition and standard CBM-I condition.
Table 3. Pearson correlation coefficients between Immersion Experience questionnaire and Slater-Usoh-Steed questionnaire scores and changes in state anxiety and in anxiety reactivity due to the stressor.

<table>
<thead>
<tr>
<th>Emotional Outcomes</th>
<th>IEQa challenge</th>
<th>IEQ control</th>
<th>IEQ dislocation</th>
<th>IEQ emotional involvement</th>
<th>IEQ cognitive involvement</th>
<th>SUSb</th>
</tr>
</thead>
<tbody>
<tr>
<td>State anxiety change</td>
<td>−0.10</td>
<td>−0.49c</td>
<td>−0.51c</td>
<td>−0.33d</td>
<td>−0.52e</td>
<td>−0.18</td>
</tr>
<tr>
<td>Anxiety reactivity to stressor</td>
<td>−0.16</td>
<td>−0.26</td>
<td>−0.36d</td>
<td>−0.17</td>
<td>−0.31d</td>
<td>−0.12</td>
</tr>
</tbody>
</table>

aIEQ: Immersion Experience questionnaire.
bSUS: Slater-Usoh-Steed questionnaire.
cP<.01.
dP<.05.
eP<.001.

We further computed the correlation between changes in interpretation bias scores (amount of positive target interpretations and amount of negative target interpretations) as a result of the training intervention (difference in interpretation bias score between post-training and pretraining) with changes in state anxiety and anxiety reactivity and with the IEQ and SUS scores. None of the correlations were significant (Pearson r range=[−0.13 to 0.20], P>.05).

Discussion

Principal Findings

In this proof-of-principle study, we examined the use of stereoscopic 3D VR technology to enrich the training experience and ultimately enhance the effects of CBM-I training for performance anxiety. The main idea behind the study was that embedding the training scenarios in a virtual environment where participants could immerse themselves and explore would improve the participants’ engagement with the training and amplify the activation of (anxiety-relevant) schemas and the related individual’s habitual pattern of biased information processing (ie, negative interpretation bias), which are necessary ingredients for this type of intervention to succeed.

When examining participants’ experience with the training, the VR-CBM-I group experienced a higher degree of immersion and presence during the training than the standard CBM-I group. In particular, the results showed that there was a significantly higher level of perceived control, real-world dissociation, and emotional and cognitive involvement for participants in the VR-CBM-I group, whereas there was no significant group difference in the level of perceived challenge. Consistent with previous studies where standard CBM-I interventions have shown to reduce state anxiety levels [11-13], all participants showed an overall decline in state anxiety after the training. As hypothesized, these reductions were significantly more pronounced in the VR-CBM-I group compared with the standard CBM-I. In addition, lower anxiety reactivity to a stressor was observed in the VR-CBM-I compared with the standard CBM-I. Furthermore, a greater feeling of emotional involvement and being in control within the scenarios were also positively associated with reductions in state anxiety. Conversely, greater feelings of presence were not associated with any change in state anxiety or emotional reactivity.

Posthoc analyses showed that a higher degree of cognitive involvement in the training scenarios and a greater perception of dissociation from the outside real world were related to both a greater reduction in state anxiety and lower anxiety reactivity to the stressor. Furthermore, a greater feeling of emotional involvement and being in control within the scenarios were also positively associated with reductions in state anxiety. Altogether, the results of the study seem to suggest a combination of specific and nonspecific effects of the VR-based CBM-I training on anxiety. The 2 versions of the training did not differ in the successful manipulation of the targeted interpretation bias for threatening information: all participants showed a decrease in the tendency to interpret ambiguous information negatively in favor of more benign interpretations. Furthermore, although both groups showed a decrease in state anxiety, VR-CBM-I training induced a steeper reduction in state anxiety and a blunted emotional response to the stressor. Supposedly, the combination of the CBM-I training mechanisms and other VR-specific factors may have enhanced these effects. Although to be taken cautiously, the positive correlations between changes in state anxiety and anxiety stress reactivity and the control, cognitive and emotional involvement, and real-world dissociation components of the immersive experience in the virtual environment seem to support this hypothesis. By experiencing the scenarios in a deeper fashion—hence, by more effectively activating the biased threat-related interpretive schemata—the training effects on basic information cognitive processing would more easily generalize to stronger emotional effects, as observed in the VR-CBM-I group.

Limitations

Despite the very promising results, no definite conclusion on the (clinical) effectiveness of VR-CBM-I can currently be drawn. Being the very first combining VR and CBM-I, this study was primarily concerned with examining the feasibility and potential of VR-CBM-I training, by focusing, as a first step, on comparing the delivery modes of the training within a semiexperimental design. Therefore, the lack of a full control condition (ie, a placebo or neutral CBM-I training group) prevents from claiming that VR-CBM-I is more effective than the standard CBM-I. The next step in the evaluation of VR-CBM-I would consequently involve a full factorial experimental design, combining the 2 delivery modalities (VR
yes vs no) and the 2 intervention components (active vs neutral CBM-I), to (1) experimentally compare the effects of both interventions against a neutral condition with no active training ingredients and (2) disentangle the active effects of the VR environment from the CBM-I training specific effects.

Relatedly, according to the preliminary phase of the study, participants completed only 1 session of training in the lab. Although the VR-based CBM-I successfully impacted on emotional outcomes in the immediate term and in response to a stressor, the duration of the effects over time is yet to be tested against a full control condition, as well as the exposure to multiple sessions of training over time. These latter aspects are particularly crucial in the view of effectively deploying (mobile) VR-based CBM interventions. The findings of this study are also promising regarding the boredom participants experience with multiple sessions of standard CBM-I [18,19].

The results of our study are restricted to the type of anxiety considered (ie, performance anxiety) and the self-selected group of undergraduate university students based on convenience sampling. Although students actively responded to flyers advertising the training as a tool they could use to do something for their test stress and anxiety, they were all compensated for participation, which might have involved an exaggeration of their initial levels of trait anxiety to be included in the study. Relatedly, the preliminary nature of the study involved recruiting only a small sample, resulting in an overall lack of power for the generalization of the behavior-change effects of the VR-CBM-I. Whether the results of this study may be generalized to other forms of (more severe) anxiety and groups of patients will need to be further investigated in a larger study with a self-motivated target population (eg, patients with anxiety problems).

Finally, the study points us to a number of key design questions. First, within the scope of this experiment, it is not yet clear how or to what extent the various perceptual factors within the 3D virtual environment (eg, the 3D background view, ambient noises, animation, and blur) influenced the outcomes of the VR-CBM-I training. From a design perspective, the deployment of highly controlled and more sophisticated experimental designs would allow us to achieve a greater insight to further optimize the mobile VR training intervention, allowing us to isolate and compare the effects of different technical features on the users’ perception of the virtual environment and the working mechanisms of the intervention (eg, trials of intervention principles [44]). For example, in this study, the training scenarios were embedded in the corresponding virtual environment as pop-up text appearing on the user’s visual field, which could be perceived as being artificial or not realistic enough. The use of audio narration of the scenarios may be a feasible option in the future development of the VR system to enhance both the training experience and the activation of the targeted emotional response [45,46]. Furthermore, given that the user interactions within the current mobile VR system were restricted to the presentation of premade scenarios and situations, future VR-CBM-I developments could investigate the use of a more interactive mobile VR system, allowing the scenarios, situations, and environments to unfold based on the choices and actions of the users. This could potentially afford a larger degree of freedom to explore and interact with the computer-generated virtual space, which could more effectively mirror users’ (emotional) experience and interaction with their real daily environment.

**Conclusions**

To conclude, this proof-of-principle study is the first investigating the feasibility and potential of using mobile VR technology to deliver CBM-I training for anxiety problems. When compared with the standard CBM-I training, a mobile VR-based CBM-I training improved the users’ experience with the training program and produced greater beneficial effects on anxiety-related emotional outcomes, while similarly changing the targeted cognitive processes. This study provided first evidence that (1) the putative working principles underlying CBM-I trainings can be translated into a virtual environment, and (2) stereoscopic 3D mobile VR technology appears to be a promising technological affordance to boost the effects of such a class of interventions, while increasing users’ experience with the training application.

**Acknowledgments**

This study was supported by the Department of Engineering and Digital Arts of the University of Kent, Kent, United Kingdom.

**Conflicts of Interest**

None declared.

**Multimedia Appendix 1**

Video showing some examples of the virtual environments included in the virtual reality Cognitive Bias Modification of Interpretations training intervention.

[MP4 File (MP4 Video), 19MB - mental_v6i2e11517_app1.mp4 ]

**References**


36. Julian LJ. Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). Arthritis Care Res (Hoboken) 2011 Nov;63(Suppl 11):S467-S472 [FREE Full text] [doi: 10.1002/acr.20561] [Medline: 22588767]


Abbreviations

ANOVA: analysis of variance
CBM: Cognitive Bias Modification
CBM-I: Cognitive Bias Modification of Interpretations
drm: decreased measures
IEQ: Immersion Experience Questionnaire
STAI: State-Trait Anxiety
SUS: Slater-Usoh-Steed
VAS: visual analogue scales
VR: virtual reality
VR-CBM-I: virtual reality Cognitive Bias Modification of Interpretations
3D: three-dimensional

Edited by G Eysenbach; submitted 10.07.18; peer-reviewed by D Leightley, D Carvalho, L Notebaert; comments to author 07.10.18; revised version received 07.11.18; accepted 10.11.18; published 21.02.19.

Please cite as:
Believing Is Seeing: A Proof-of-Concept Semiexperimental Study on Using Mobile Virtual Reality to Boost the Effects of Interpretation Bias Modification for Anxiety
JMIR Ment Health 2019;6(2):e11517
URL: http://mental.jmir.org/2019/2/e11517/
doi:10.2196/11517
PMID:30789353

©Boris Otkhmezuri, Marilisa Boffo, Panote Siriaraya, Maria Matsangidou, Reinout W Wiers, Bundy Mackintosh, Chee Siang Ang, Elske Salemink. Originally published in JMIR Mental Health (http://mental.jmir.org), 21.02.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.
Identifying Behaviors Predicting Early Morning Emotions by Observing Permanent Supportive Housing Residents: An Ecological Momentary Assessment

Rajesh R Nandy¹, BStat, MStat, PhD; Karabi Nandy¹, PhD; Emily T Hébert², DrPH; Michael S Businelle³, PhD; Scott T Walters⁴, PhD

¹Department of Biostatistics and Epidemiology, School of Public Health, University of North Texas Health Science Center, Fort Worth, TX, United States
²Oklahoma Tobacco Research Center, The University of Oklahoma Health Sciences Center, Oklahoma City, OK, United States
³Department of Family and Preventive Medicine, The University of Oklahoma Health Sciences Center, Oklahoma City, OK, United States
⁴Department of Health Behavior and Health Systems, School of Public Health, University of North Texas Health Science Center, Fort Worth, TX, United States

Corresponding Author:
Rajesh R Nandy, BStat, MStat, PhD
Department of Biostatistics and Epidemiology
School of Public Health
University of North Texas Health Science Center
3500 Camp Bowie Boulevard
Fort Worth, TX, 76107
United States
Phone: 1 2063296441
Email: Rajesh.Nandy@unthsc.edu

Abstract

Background: Behavior and emotions are closely intertwined. The relationship between behavior and emotions might be particularly important in populations of underserved people, such as people with physical or mental health issues. We used ecological momentary assessment (EMA) to examine the relationship between emotional state and other characteristics among people with a history of chronic homelessness who were participating in a health coaching program.

Objective: The goal of this study was to identify relationships between daily emotional states (valence and arousal) shortly after waking and behavioral variables such as physical activity, diet, social interaction, medication compliance, and tobacco usage the prior day, controlling for demographic characteristics.

Methods: Participants in m.chat, a technology-assisted health coaching program, were recruited from housing agencies in Fort Worth, Texas, United States. All participants had a history of chronic homelessness and reported at least one mental health condition. We asked a subset of participants to complete daily EMAs of emotions and other behaviors. From the circumplex model of affect, the EMA included 9 questions related to the current emotional state of the participant (happy, frustrated, sad, worried, restless, excited, calm, bored, and sluggish). The responses were used to calculate two composite scores for valence and arousal.

Results: Nonwhites reported higher scores for both valence and arousal, but not at a statistically significant level after correcting for multiple testing. Among momentary predictors, greater time spent in one-on-one interactions, greater time spent in physical activities, a greater number of servings of fruits and vegetables, greater time spent interacting in a one-on-one setting as well as adherence to prescribed medication the previous day were generally associated with higher scores for both valence and arousal, and statistical significance was achieved in most cases. Number of cigarettes smoked the previous day was generally associated with lower scores on both valence and arousal, although statistical significance was achieved for valence only when correcting for multiple testing.

Conclusions: This study provides an important glimpse into factors that predict morning emotions among people with mental health issues and a history of chronic homelessness. Behaviors considered to be positive (eg, physical activity and consumption of fruits and vegetables) generally enhanced positive affect and restrained negative affect the following morning. The opposite was true for behaviors such as smoking, which are considered to be negative.
permanent supportive housing; circumplex model of affect; ecological momentary assessment; emotion; valence; arousal; hierarchical mixed effects model; mobile phone

Introduction

More than half a million individuals are homeless at any given time in the United States [1]. Homelessness is associated with a higher prevalence of mental illness, higher rates of morbidity and mortality, and increased rates of drug abuse, criminality, and victimization [2]. Permanent supportive housing (PSH) is one approach to reducing chronic homelessness and provides low-cost community-based housing alongside supportive services. PSH has been demonstrated to reduce homelessness, increase housing tenure, and decrease emergency room visits and hospitalization [3]. Although PSH can result in lower overall costs to society, people who reside in PSH face numerous challenges in their ability to live independently, including, in many cases, physical and mental health conditions requiring treatment. For example, 73% of PSH residents in Fort Worth, Texas, United States, reported at least one chronic health condition, 55% reported having received treatment for a mental health condition, 67% reported having a history of substance abuse, and 44% reported both co-occurring substance abuse and mental health concerns [4].

Mood and emotional reactivity play an important role in both mental and physical health. For example, Gallo and Matthews [5] found that negative emotions and cognitions were related to cardiovascular disease and all-cause mortality and contributed to the relationship between socioeconomic status and health [6]. A study of well-being among adults in England associated positive affect with survival, even after controlling for demographic factors and baseline health [7]. In addition, multiple studies have shown the association of anxiety, stress, and negative affect with health behaviors such as smoking, alcohol, and drug use [8-12]. Additional research is needed to examine how health behaviors affect affect and stress in disadvantaged and understudied adults.

Ecological momentary assessment (EMAs) techniques use mobile devices to assess thoughts, feelings, and behaviors in real-time in an individual’s natural setting [13]. A review of EMA studies on mood disorders and dysregulation demonstrated that real-time assessment reduces recall bias and allows for the study of dynamic processes and context-specific relationships related to mood [14]. For instance, in one 4-day long EMA study of depression among adolescents, higher pretreatment positive affect, lower negative affect, and a higher positive-to-negative affect ratio predicted a lower clinician-rated severity of problems following treatment [15]. The measures of affect were created using items adapted from the Positive and Negative Affect Scale for Children [16]. Similarly, another EMA study of affect and depressive illness found that response to treatment was predicted by daily increases in positive affect among individuals with clinical depression [17]. EMAs involving substance use in adolescents indicate that alcohol intake and cigarette intake are predicted by greater negative mood states, including sadness, depression, anger, and stress, as well as greater conduct and behavioral problems [18-21]. Overall, EMAs may be a useful way to monitor and, ultimately, intervene to prevent maladaptive mood experience and mood regulation processes [22].

Although EMA has been used to evaluate dynamic changes in mood and behavior, no study to date has examined the relationship between emotions and behavior among adults in PSH. The purpose of this study was to explore the prospective associations between emotions (ie, valence and arousal) and health behaviors among adults residing in PSH using EMA. Considering the relatively high costs associated with physical and mental health disorders in this population, it can be beneficial to identify factors affecting daily emotion patterns in order to predict and intervene with persons who are at risk.

Methods

Participants, Design, and Study Procedures

We obtained data for this study from the Mobile Community Health Assistance for Tenants (m.chat) project, a technology-assisted health coaching intervention designed to improve health indicators among PSH residents in Fort Worth, Texas [23]. We recruited participants via convenience sampling from 6 local housing agencies in Fort Worth. The participants were adult, English speaking, Medicaid-enrolled or eligible, and reported at least one of the following conditions in the past year: prescribed medication for psychological or emotional problems, experienced hallucinations, received a pension for a psychiatric disability, or reported at least moderate levels of depression (>9 on the Patient Health Questionnaire). Participants met monthly with a health coach who helped to set goals related to diet, exercise, substance use, medication compliance, social support, and recreation or leisure. We gave the opportunity to participate in the EMA portion of the project to a subgroup of participants who scored ≥4 on the Rapid Estimate of Adult Literacy in Medicine-Short Form (indicating >6th grade English literacy level). This subset of clients completed EMAs each morning with questions about current emotions and setting as well as health behaviors from the previous day, including diet, exercise, substance use, leisure time activities, medication compliance, and social interactions. We used an EMA protocol based on that of a previous study with homeless smokers [20]. We provided participants with a smartphone and granted unlimited voice, short message service text, and 2 GB data for their personal use. While enrolled in the EMA portion of m.chat, participants received up to 15 “Chat Bucks” each month, proportional to the percentage of days they completed the assessment (1 Chat Buck=US $1 redeemable for health-related supplies; thus, participants could earn up to US $15 worth of health-related supplies each month). Provided they were compliant with at least 50% of the EMA prompts, participants could carry the phone for up to 12 months. Project resources allowed for up to 80 participants to participate in the EMA.
portion at any one time; when participants returned the phone (because they had reached the end of their allotted time, were failing to complete assessments, or decided they did not want to carry the phone any longer), it was reset to factory settings and offered to another participant based on the order of enrollment into the parent study.

The Institutional Review Board of the University of North Texas Health Science Center approved this project, and we assured participants of confidentiality. All participants gave informed consent.

For analysis, we included 155 participants who completed a total of 18,357 daily assessments between May 1, 2016, and April 30, 2017. On average, individuals received 139 daily assessments or prompts (range 14-334) and completed 106 assessments (range 4-322). The sample was split almost evenly between males (n=77) and females (n=78), and the average age was 52 (SD 8) years.

**Instruments and Measures**

The mobile app alerted participants to complete an assessment 30 minutes after the participant’s self-reported waking time. We asked the participants to complete the assessment within 30 minutes of the initial alert; they had the option to “snooze” an assessment request 3 times each day before the EMA would be counted as missed. Below are the questions that were presented in the daily EMA (we have only presented the questions or response options considered in the analyses).

**Emotions**

We measured 9 emotions items on a Likert-type scale from 1 (strongly disagree) to 5 (strongly agree): I feel happy, I feel frustrated, I feel sad, I feel worried, I feel restless, I feel excited, I feel calm, I feel bored, and I feel sluggish.

**Physical Activity**

Participants were asked how many hours they spent sitting, how many minutes they walked or biked to get somewhere, how many minutes they were physically active for fitness (eg, running or sports), and how many minutes they were physically active at work or home (eg, cleaning, lifting, or carrying things) the previous day.

**Diet**

We asked participants how many servings of fruits and vegetables they ate, how many sugar-sweetened beverages they drank, and how many desserts and other sweets they ate the previous day.

**Social Support**

We asked the participants about total minutes they spent in meaningful one-on-one conversations with other people and the total minutes they spent in meaningful group interactions (eg, going to church, participating in an exercise class, or other social occasions) the previous day.

**Medication Compliance**

We asked participants whether they took all of their medication as prescribed the previous day.

**Tobacco Use**

We asked participants whether they used tobacco (cigarettes) the previous day, and if so, how many cigarettes they smoked.

Demographic characteristics such as age, sex, and race (white or nonwhite), collected at baseline, were used as covariates in the analyses.

**Statistical Modeling and Analysis**

The circumplex model of affect [24,25] was used to categorize each emotion in a 2-dimensional circular space, containing dimensions for arousal and valence (Figure 1). Valence, in the context of emotions, is defined as the intrinsic attractiveness or averseness of an event, object, or situation. Likewise, arousal is the state of being physiologically and mentally alert, awake, and attentive. In a recent refinement of the model using regression [26], the circumplex model was quantitatively visualized as a circular space of radius of 1 unit within a 2-dimensional Cartesian coordinate system, which assigns scores for valence and arousal for each emotion. This model includes a comprehensive list of mood items commonly considered in behavioral sciences. We obtained scores for valence and arousal in the circumplex from this model for each of the 9 emotion items considered (Table 1). In Figure 2, these 9 emotions are presented within the circumplex, depicting their valence and arousal coordinates. During each daily assessment, we created composite scores of valence and arousal as weighted sums of responses from all 9 emotion questionnaire items, with the reported emotion scores serving as weights. These two composite scores were the outcomes of this study. Specifically, let a subject’s response in Likert-type scale for I feel happy, I feel frustrated, I feel sad, I feel worried, I feel restless, I feel excited, I feel calm, I feel bored, and I feel sluggish be denoted by l 1, l 2, l 3, l 4, l 5, l 6, l 7, l 8, and l 9, respectively. Then, using Table 1, the composite score for valence is 0.95x l 1-0.5x l 2-0.95x l 3-0.15x l 4-0.15x l 5+0.7x l 6+0.75x l 7-0.4x l 8-0.15x l 9. The corresponding composite score for arousal is 0.15x l 1+0.4x l 2-0.4x l 3-0.3x l 4+0.3x l 5+0.7x l 6-0.7x l 7-0.8x l 8-0.5x l 9. In Table 2, for each emotion, we have provided the mean of the subject means, SD of the subject means (between subjects), and mean of the subject SDs (within subjects). All questions in the various domains (eg, diet and physical activity) described in the Instruments and Measures section, except the emotions items, were considered as potential predictors of the outcomes. To reduce the large number of predictors in the model, we combined some of the variables within the same domain to create the following new variables as predictors: number of servings of healthy diet, number of servings of sweets, and number of minutes of total physical activity. In Table 3, for each momentary predictor, we have provided the mean of the subject means, SD of the subject means (between subjects), and mean of the subject SDs (within subjects). It is important to observe that even though the EMA emotions questions asked about present emotions, predictors were recalled values from the previous day.
Figure 1. Circumplex model of affect.

Table 1. Circumplex scores for the emotions considered.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Valence</th>
<th>Arousal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>0.95</td>
<td>0.15</td>
</tr>
<tr>
<td>Frustrated</td>
<td>-0.50</td>
<td>0.40</td>
</tr>
<tr>
<td>Sad</td>
<td>-0.95</td>
<td>-0.40</td>
</tr>
<tr>
<td>Worried</td>
<td>-0.15</td>
<td>-0.30</td>
</tr>
<tr>
<td>Restless</td>
<td>-0.15</td>
<td>0.30</td>
</tr>
<tr>
<td>Excited</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Calm</td>
<td>0.75</td>
<td>-0.70</td>
</tr>
<tr>
<td>Bored</td>
<td>-0.40</td>
<td>-0.80</td>
</tr>
<tr>
<td>Sluggish</td>
<td>-0.15</td>
<td>-0.50</td>
</tr>
</tbody>
</table>
Figure 2. Circumplex model in Cartesian coordinate system.

Table 2. Descriptive statistics for the 9 emotion outcomes.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Between-subject SD</th>
<th>Mean within-subject SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>3.54</td>
<td>0.74</td>
<td>0.70</td>
</tr>
<tr>
<td>Sad</td>
<td>2.55</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Restless</td>
<td>2.59</td>
<td>0.83</td>
<td>0.70</td>
</tr>
<tr>
<td>Excited</td>
<td>3.09</td>
<td>0.78</td>
<td>0.70</td>
</tr>
<tr>
<td>Calm</td>
<td>3.42</td>
<td>0.71</td>
<td>0.67</td>
</tr>
<tr>
<td>Sluggish</td>
<td>2.71</td>
<td>0.92</td>
<td>0.74</td>
</tr>
<tr>
<td>Frustrated</td>
<td>2.55</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Worried</td>
<td>2.60</td>
<td>0.85</td>
<td>0.74</td>
</tr>
<tr>
<td>Bored</td>
<td>2.41</td>
<td>0.80</td>
<td>0.66</td>
</tr>
</tbody>
</table>
Table 3. Descriptive statistics for the quantitative momentary predictors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Between-subject SD</th>
<th>Mean within-subject SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total physical activity</td>
<td>34.62</td>
<td>24.67</td>
<td>20.78</td>
</tr>
<tr>
<td>Minutes of one-on-one interaction</td>
<td>79.97</td>
<td>49.28</td>
<td>48.94</td>
</tr>
<tr>
<td>Minutes spent in group interaction</td>
<td>44.41</td>
<td>32.74</td>
<td>36.97</td>
</tr>
<tr>
<td>Hours spent sitting</td>
<td>5.47</td>
<td>1.93</td>
<td>2.13</td>
</tr>
<tr>
<td>Fruits and vegetables</td>
<td>2.80</td>
<td>1.55</td>
<td>1.23</td>
</tr>
<tr>
<td>Sweets</td>
<td>2.60</td>
<td>1.67</td>
<td>1.13</td>
</tr>
<tr>
<td>Number of cigarettes</td>
<td>3.49</td>
<td>5.12</td>
<td>1.66</td>
</tr>
</tbody>
</table>

In that sense, the predictors are not strictly momentary, but will be referred to as momentary variables for statistical modeling and analysis. Individual demographic characteristics (ie, age at the onset of the EMA study, sex, and race) were considered time invariant for the duration of the study. Race was dichotomized as white and nonwhite, as 94.5% (147/155) of participants were either white or African American individuals.

In the general statistical model for the analysis, for each outcome, we denote the response on the \( t \)th assessment from the \( i \)th subject by \( Y_{it} \), the value of the \( j \)th demographic predictor (out of \( k \) total number of predictors) from the \( i \)th subject by \( X_{ij} \) and the value of the momentary predictor on the \( t \)th assessment from the \( i \)th subject by \( Z_{it} \). Then the hierarchical model can be presented as follows:

\[
Y_{it} = \beta_0 + \beta_1 X_{ij} + \beta_2 Z_{it} + \epsilon_{it}
\]

All analyses were performed using MIXED procedure in SAS (SAS Institute) with the intercept specified as a random effect and within-subject residuals specified to have a first-order autoregressive correlation.

Since there are eleven predictors in our model, we implemented the popular Bonferroni correction to adjust the reported \( P \) values for the predictors. It should be noted that the very conservative Bonferroni-corrected \( P \) value threshold of .05 is equivalent to an unadjusted \( P \) value threshold of .0045. Since we were not interested in the statistical significance of the intercept term, we did not consider it for the Bonferroni correction.

Results

Associations Among Momentary Variables and Emotions Controlling for Demographic Characteristics

Analyses of the associations between momentary variables and valence and arousal were performed, controlling for the 3 demographic predictors of age, sex, and race. The results for the valence and arousal outcomes are presented in Tables 4 and 5, respectively (all the \( P \) values correspond to 2-tailed tests). For the predictors with an unadjusted \( P \) value<.05, we also presented Bonferroni-corrected \( P \) values (in parentheses). None of the demographic variables predicted either outcome at a statistically significant level after the extreme Bonferroni adjustment. However, the effect of race on valence barely missed significance after adjustment (unadjusted \( P \) value=.007; adjusted \( P \) value=.08), with white individuals reporting much lower valence scores on average.

Minutes spent doing physical activity the previous day was a statistically significant predictor of both valence and arousal, with expected higher scores for increased physical activity. Time spent in meaningful group interaction the previous day was not a statistically significant predictor of either valence or arousal. Time spent in meaningful one-on-one social interaction the previous day was a statistically significant predictor of both valence and arousal, with expected higher scores for more interaction time. Hours spent sitting the previous day was a statistically significant predictor of both valence and arousal, with an expected lower score for an increase in time spent sitting. Number of total servings of fruits and vegetables consumed the previous day was a statistically significant predictor of both valence and arousal, with expected higher scores for greater servings. Number of total servings of sugar-sweetened beverages and desserts the previous day was not a statistically significant predictor of either valence or arousal. Adherence to medication the previous day was a statistically significant predictor of both valence and arousal, with higher scores for adherence. Any tobacco usage the previous day was a statistically significant predictor of only valence; on average, smoking a higher number of cigarettes resulted in lower valence scores.
### Table 4. Results for valence with momentary predictors, controlling for demographic characteristics.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Estimate</th>
<th>SE</th>
<th>t (df)(^a)</th>
<th>P value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.32</td>
<td>1.422</td>
<td>1.63 (152)</td>
<td>.10</td>
</tr>
<tr>
<td>Age</td>
<td>-0.008</td>
<td>0.027</td>
<td>-0.28 (153)</td>
<td>.78</td>
</tr>
<tr>
<td>Male</td>
<td>-0.49</td>
<td>0.459</td>
<td>-1.07 (153)</td>
<td>.28</td>
</tr>
<tr>
<td>Caucasian</td>
<td>-1.23</td>
<td>0.457</td>
<td>-2.69 (153)</td>
<td>.007 (.08)</td>
</tr>
<tr>
<td>Total physical activity</td>
<td>0.007</td>
<td>0.0008</td>
<td>8.80 (153)</td>
<td>&lt;.001 (&lt;.001)</td>
</tr>
<tr>
<td>Minutes of one-on-one interaction</td>
<td>0.004</td>
<td>0.0004</td>
<td>9.53 (153)</td>
<td>&lt;.001 (&lt;.001)</td>
</tr>
<tr>
<td>Minutes spent in group interaction</td>
<td>0.00008</td>
<td>0.0005</td>
<td>0.17 (153)</td>
<td>.87</td>
</tr>
<tr>
<td>Hours spent sitting</td>
<td>-0.04</td>
<td>0.011</td>
<td>-3.42 (153)</td>
<td>.001 (.007)</td>
</tr>
<tr>
<td>Fruits and vegetables</td>
<td>0.12</td>
<td>0.014</td>
<td>8.54 (153)</td>
<td>&lt;.001 (&lt;.001)</td>
</tr>
<tr>
<td>Sweets</td>
<td>0.02</td>
<td>0.016</td>
<td>1.34 (153)</td>
<td>.18</td>
</tr>
<tr>
<td>Medication</td>
<td>0.76</td>
<td>0.105</td>
<td>7.23 (153)</td>
<td>&lt;.001 (&lt;.001)</td>
</tr>
<tr>
<td>Number of cigarettes</td>
<td>-0.06</td>
<td>0.008</td>
<td>-7.99 (153)</td>
<td>&lt;.001 (&lt;.001)</td>
</tr>
</tbody>
</table>

\(^a\)All the P values correspond to 2-tailed tests.
\(^b\)Bonferroni-corrected P values are in parentheses.

### Table 5. Results for arousal with momentary predictors and controlling for demographic characteristics.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Estimate</th>
<th>SE</th>
<th>t (df)(^a)</th>
<th>P value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.97</td>
<td>0.520</td>
<td>-5.72 (152)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0001</td>
<td>0.010</td>
<td>-0.01 (153)</td>
<td>.99</td>
</tr>
<tr>
<td>Male</td>
<td>-0.17</td>
<td>0.167</td>
<td>-1.04 (153)</td>
<td>.30</td>
</tr>
<tr>
<td>Caucasian</td>
<td>-0.42</td>
<td>0.167</td>
<td>-2.50 (153)</td>
<td>.01 (.13)</td>
</tr>
<tr>
<td>Total physical activity</td>
<td>0.003</td>
<td>0.0004</td>
<td>7.40 (153)</td>
<td>&lt;.0001 (&lt;.001)</td>
</tr>
<tr>
<td>Minutes of one-on-one interaction</td>
<td>0.002</td>
<td>0.0002</td>
<td>11.17 (153)</td>
<td>&lt;.0001 (&lt;.001)</td>
</tr>
<tr>
<td>Minutes spent in group interaction</td>
<td>0.0003</td>
<td>0.0002</td>
<td>1.67 (153)</td>
<td>.09</td>
</tr>
<tr>
<td>Hours spent sitting</td>
<td>-0.02</td>
<td>0.005</td>
<td>-4.20 (153)</td>
<td>&lt;.001 (&lt;.001)</td>
</tr>
<tr>
<td>Fruits and vegetables</td>
<td>0.03</td>
<td>0.006</td>
<td>4.07 (153)</td>
<td>&lt;.001 (&lt;.001)</td>
</tr>
<tr>
<td>Sweets</td>
<td>-0.009</td>
<td>0.007</td>
<td>-1.36 (153)</td>
<td>.17</td>
</tr>
<tr>
<td>Medication</td>
<td>0.15</td>
<td>0.045</td>
<td>3.26 (153)</td>
<td>.001 (.01)</td>
</tr>
<tr>
<td>Number of cigarettes</td>
<td>-0.006</td>
<td>0.003</td>
<td>-1.88 (153)</td>
<td>.06</td>
</tr>
</tbody>
</table>

\(^a\)All the P values correspond to 2-tailed tests.
\(^b\)Bonferroni-corrected P values are in parentheses.

### Influence of Demographic Variables on Momentary Predictors for Emotions

Even though the model controls for demographic characteristics in analyzing the effect of momentary variables on valence and arousal, it is worthwhile to explore how much influence the demographic predictors have on the momentary predictors. A strong influence of demographic predictors on the momentary predictors can make the regression coefficients unstable and hard to interpret. Unlike in a standard multiple regression framework, in our hierarchical model, the influence cannot be measured directly by studying the multicollinearity properties and other standard regression diagnostics. Instead, the amount of influence can be indirectly measured by analyzing two additional models: one with only demographic predictors and one with only momentary predictors. The change in values of the estimated regression coefficients in the full models compared with the two isolated models described above can be used to assess the influence and the robustness of the coefficients.

For the sake of brevity, we did not present the actual results from the two isolated models here, but the results are remarkably consistent with our findings from the combined model in the previous section. Not only do the statistical significances of the momentary predictors match closely but also the individual estimates of the regression coefficients are surprisingly close. The individual estimates of the regression coefficients are very close for the demographic predictors as well. The observed
consistency provides fairly strong evidence on the orthogonality of the demographic predictors from the momentary predictors.

**Discussion**

**Principal Findings**

These findings provide an important glimpse into factors that affect valence and arousal in a population of individuals residing in supportive housing. To our knowledge, this is the first study to examine the connection between emotions and other factors among people with mental health disorders and a history of chronic homelessness. This underserved population is often excluded from research studies due to co-occurring mental and physical disorders, resulting in substantial gaps in our understanding of their health and health behaviors.

Our analyses provide a number of observations about the relationships between health behaviors and subsequent emotions. First, we found that physical activity was significantly associated with positive emotions the following day. This finding is consistent with the literature showing the association of moderate physical activity with improved and maintained mood [27] as well as decreased symptoms of depression and anxiety [28]. We also found a positive relationship between fruit and vegetable intake and emotions the following day. Similarly, a study using data from the Canadian Community Health Survey found a significant association of fruit and vegetable intake with lower odds of depression and psychological distress [29]. Taken together, these results argue that a coordinated program to improve physical activity and diet should be a fundamental part of health interventions for people with mental health disorders and a history of chronic homelessness.

We also found that smoking cigarettes had a negative effect on valence the following day. Although nicotine may have a calming effect due to the inhibition of negative emotions such as anger [30], our results suggest that this effect may not carry forward to the following day. Research has suggested that nicotine dependency exacerbates stress [31], and a meta-analysis of changes in mental health after smoking cessation revealed that smoking cessation is associated with reduced depression, anxiety, and stress, with effect sizes equal to or larger than those of antidepressant treatment for mood disorders [32]. Thus, individuals receiving treatment for mood disorders may benefit from concurrent smoking cessation therapy.

We also found a strong relationship between the amount of time spent in individual social interactions and emotions. Interestingly, “time spent on meaningful one-on-one social interaction the previous day” was strongly associated with arousal and valence, while the “amount of time spent interacting in a group setting” was not significantly associated with emotions. This finding was unexpected given the substantial evidence that social support predicts the quality of life in many areas [33]. However, supportive housing residents are encouraged to attend support groups that address lifestyle skills, chronic disease management, and substance use. For this population, it may be that group interactions do not contribute to emotions unless the individual feels personally connected to at least one other person in the group. Thus, group interactions by themselves may not predict emotions, while individual interactions outside of the group setting may be one indicator of healthy, rewarding relationships.

In the analysis of the effect of demographics on the association between momentary predictors with valence or arousal, demographic variables had minimal effects on the regression coefficients of the momentary predictors, even when statistically significant. Hence, it is reasonable to conclude that the demographic predictors operated almost independently of the momentary variables in terms of influencing emotions.

Finally, it would be possible to study the association of emotional affect with subsequent same-day behaviors, for instance, examining the effects of emotions in the morning on smoking or drinking later in the day. Such analyses are beyond the scope of this paper, but we plan to investigate these associations in a future manuscript.

**Limitations and Strengths**

Our study had a number of limitations. Notably, our protocol included only daily morning assessments. Thus, we were not able to examine within-day variability. However, unlike other EMA studies, which typically run for a few days to, at most, a few weeks, our study ran up to 334 days with an average of 156 days of monitoring among all participants. This allowed us to examine associations for a much longer period than most other EMA studies. In addition, our results are generalizable only to a population of individuals residing in PSH with a history of homelessness and mental health issues. It is unclear whether the findings are generalizable to other groups of people with mental health problems, let alone the population in general. Relatedly, all participants self-reported depression or a mental health condition at baseline, which may have affected emotions and mood independently of other behavioral measures. For instance, the average client reported a score of 12.62 on the Patient Health Questionnaire, indicating that most clients felt at least moderate levels of depression upon admission to the program. Finally, we cannot rule out the possibility that participating in the coaching intervention affected the relationship between behaviors and emotion. Our results must be interpreted in the context of the larger services that people were receiving in this program. Further study with a more diverse population is necessary to make any broader assertions.

**Conclusions**

Despite the limitations, our study offers an important glimpse into health behaviors that affect daily emotional arousal and valence of persons with a history of chronic homelessness and mental health problems. One of the goals of the m.chat program was to provide individual support and assistance in meeting health goals. Because mood was an important target of the program, identifying factors that predicted positive affect can help improve future iterations of programs like this. To that end, identifying modifiable behaviors associated with negative and positive moods is a first step toward improving stability and preventing future homelessness. Understanding factors associated with mood and behaviors, particularly in vulnerable populations such as formerly homeless individuals, can also...
help providers design more targeted treatment plans and provide more appropriate referrals to ancillary care services [34].

Notably, many of our findings are consistent with “common wisdom” drawn from other populations. Behaviors generally considered to be positive (eg, physical activities, consumption of fruits and vegetables, adherence to prescribed medication, and one-on-one social interaction) tended to enhance positive affect and restrain negative affect. The opposite was true for behaviors considered to be negative, such as smoking. In fact, the positive and negative impacts on physical health for most of these behaviors are well established, and it is noteworthy that their effects on positive and negative affect appear to be consistent with previous literature with other populations. In a separate analysis of m.chat program data, Holmes et al [35] found that participants who consumed the least Western-style foods (eg, fast food, sugar-sweetened beverages, and processed meat) had significantly lower depressive symptoms over 1 year than those who consumed the most Western-style foods. However, Holmes et al also reported no significant association between depressive symptoms and nutritious food intake and physical activity over 18 months. Because we found a relationship between fruits and vegetables, exercise, and next day mood (but not sugar and mood), the relationships between these variables may be somewhat different in the short versus long term. Such behaviors can be targeted with a goal of enhancing positive affect and restraining negative affect in order to improve the overall mental health of individuals. Such an intervention has the potential to improve treatments for individuals with mood and other psychological disorders.

Acknowledgments

Funding for this study was provided through a Medicaid 1115 Waiver to the State of Texas. Centers for Medicare and Medicaid Services had no role in the study design, collection, analysis, or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Conflicts of Interest

MSB was a paid consultant on the parent grant for this study.

References


27. Peluso MAM, Guerra de Andrade LHS. Physical activity and mental health: the association between exercise and mood. Clinics (Sao Paulo) 2005 Feb;60(2):159-164 [FREE Full text] [Medline: 15838583]


Abbreviations

EMA: ecological momentary assessment
PSH: permanent supportive housing