Original Paper

Julio Vega, Pittsburgh University, USA
Beth Bell, University of York, UK
Caitlin Taylor, York St John University, UK
Jue Xie, Monash University, Australia
Heidi Ng, Monash University, Australia
Mahsa Honary, Lancaster University, UK
Roisin McNaney, Monash University, Australia

Detecting Mental Health Behaviours Using Mobile Interactions (DeMMI): an Exploratory Study Focusing on Binge Eating

Abstract

Background: Binge eating is a subjective loss of control while eating, leading to the consumption of large amounts of food. It can cause significant emotional distress and is often accompanied by purging behaviours (eg, meal skipping, over-exercising or vomiting).

Objective: The aim of this study was to explore the potential for mobile sensing to detect indicators for binge eating episodes, with a view toward informing the design of future context-aware mobile interventions.

Methods: Our study was conducted in two stages. The first involved the development of the DeMMI app. As part of this, we conducted a consultation session to explore whether the types of sensor data we were proposing to capture were seen to be useful and appropriate, as well as gathering feedback on some specific app features relating to self-report. The second stage involved carrying out a 6-week period of data collection with 10 participants experiencing binge eating (logging both their mood and episodes of binge eating) and 10 comparison participants (logging only mood). An optional interview was conducted post-study discussing their experience with using the app, 8 participants (3 binge eating and 5 comparisons) consented.

Results: Findings showed unique differences in the types of sensor data that were triangulated with individuals' episodes (with nearby Bluetooth devices, screen and app usage features, mobility features, and mood scores showing relevance). Participants had a largely positive opinion about the app, its unobtrusive role, and its ease of use. Interacting with the app increased their awareness of and reflection around mood and their phone usage patterns. Moreover, they expressed no privacy concerns as the study information sheet alleviated these.

Conclusions: In this study, we contribute a series of recommendations for future studies wishing to scale our approach, and for the design of bespoke mobile interventions to support this population.

Keywords: Eating disorder; binge eating; mental health; mobile sensing; context-aware computing; NAP; EMA

Introduction

Binge eating is classified as a distinct period of time during which an individual experiences a subjective loss of control over eating, eating notably more or differently than usual, and feels unable to stop eating or limit the type or amount of food eaten [1]. It is thought to affect approximately 5% of females and 4% of males worldwide in some form over their life course [2]. Binge eating is a precursor for, and symptomatic of, clinical eating disorders including Bulimia Nervosa and Binge Eating Disorder [3]. In Bulimia Nervosa, binge eating is typically followed by the purging of calories from the body in an attempt to counteract a binge eating episode—for example, through vomiting, use of laxatives or diuretics, extreme dieting or excessive exercise—, whereas individuals with Binge Eating Disorder do not engage in purging practises [4]. Binge eating is often a hidden behaviour, conducted in secret [5], and can lead to extreme feelings of shame, worthlessness, and lack of control, which can have major impacts on an individual's mental health and emotional well-being [6]. Research has shown that clinical binges (ie, from those with an eating disorder diagnosis) yield the same subjective experiences before and after the binge as those reported from a non-clinical binge eating episode [7].

Anxiety and depression are prevalent in people with binge eating behaviours [8–10], and research indicates that binge eating is employed as a strategy for regulating negative affect (often used as an umbrella term to refer to emotive experiences such as mood, emotion, impulses and stress response [11]). Additionally, as people with binge eating behaviors are often within the normal to obese BMI category ranges [12], people within this population report feelings around being unworthy of mental health support and not identifying themselves as having an eating disorder. All this, paired alongside the secretive nature of binge eating behaviours, mean that the identification and treatment of binge eating is particularly challenging [5].

A growing body of literature has identified opportunities for mobile sensing (i.e. the collection and use of data collected from sensors embedded within mobile devices such as smartphones) to detect mental health behaviours such as schizophrenia [13], bipolar disorder [14] and depression [15]. With the exception of [16], who looked at heart rate variability as a risk predictor for emotional eating episodes, no research has explicitly investigated the potential for mobile sensing in relation to disordered eating behaviours. This is an underexplored area of research, which has significant opportunities for a) identifying the contextual and situational factors associated with binge eating episodes, and b) providing improved access to support and behaviour prevention through context-aware interventions [17,18].

Our research aimed to contextualise the experiences of people engaging in self-identified episodes of binge eating, with or without subsequent purging activities, by understanding the types of activities that a person might be engaging with in and around the occurrence of an episode. As this was first-of-its-kind, exploratory work, we employed a broad range of mobile sensors, which exist already in the mobile phone (e.g. location sensors can indicate if someone is spending a lot of time at home; movement sensors can give us an idea of how much activity a person has been engaging in; app usage sensors can indicate how much time someone is spending on social media or healthy eating and fitness apps). We asked participants to provide daily self-reports of their mood (collected in both the morning and evening) and to self-report any episodes of binge eating (logged through a button press to capture the time of the episode, with the option to provide further information in the form of free flowing text). This provided us with a measure of the differences in behavioural features, extracted from smartphone data, in days with and without incidents of binge eating, in an attempt to inform future context-aware mobile interventions to support this population.

We describe the development of the Detecting Mental health behaviours through Mobile Interactions (DeMMI) app, which was refined in consultation with service users. We then describe a 6-week remote study on 20 participants (10 with experiences of binge eating and 10 without any mental health issues, who reported twice daily mood logs and acted as a comparison group). Our contributions from this paper are threefold: 1) first, we provide a set of reflections around the challenges of conducting work with the binge eating population and the benefits of remote, anonymous engagement; 2) second, we provide unique insights into the successes and challenges surrounding our mobile sensing approach (from a pilot study perspective) and how this might be better scaled in the future for larger scale studies over longer periods of time; 3) finally, we provide a set of recommendations for the design of future context-aware interventions aiming to support people experiencing binge eating behaviours.

Background

Use of Ecological Momentary Assessment (EMA) for Monitoring Mood and Binge Eating

Cross-sectional and longitudinal studies are useful for understanding long-term risk factors of poor mental health, including those that contribute to binge eating (see [19], for review). However, they are much less useful for understanding the more immediate contextual and situational factors that directly contribute to fluctuations in mood that accompany binge eating behaviour. Ecological Momentary Assessment (EMA) techniques are better able to provide information about these contextual factors. EMA involves the recording of problematic mood, thoughts and behaviours, as well as the events that immediately precede them in order to identify predictive patterns. While early EMA involved the use of paper-based diaries, more recent research has made use of digital tools (eg, mobile phone and web-based apps) to gather real-time self reported data. Data collected through digitally enhanced EMA not only has the potential to enhance understanding of binge eating in research settings, but can also be used in therapeutic settings, and by individuals, to better understand and monitor individual patterns related to poor mental health. Indeed, a large proportion of smartphone apps for binge eating specifically (eg, [20–23]), and for mental health (eg, [24–30]), typically involves self-reporting and repeatedly prompting participants over time [31]. Such mobile-monitoring apps are typically well-received by young people [32].

EMA relies on the self-report of affective states, but there is much heterogeneity in the way these affective states are measured. Self-report questions used in smartphone-based EMA are often literal translations of clinical tools [24,26,33,34]. For example, The Positive and Negative Affect Scale (PANAS) involves participants indicating the extent to which they are experiencing ten types of positive (eg, "Alert") and negative (eg, "Upset") affect, using 5-point Likert scales [35]. Although the original PANAS is generally considered too long to be applied with high frequency in EMA [36], subscales and shortened forms have been delivered using smartphone EMA with good response rates, even when used multiple times daily [33]. Furthermore, there are other approaches attempting to reduce the burden on participants through the use of visual scales. For example, The Self-Assessment Manikin (SAM) uses three icon-based scales to measure pleasure, arousal, and dominance [37]. In their original form, the icons are abstract outlines of a human-like figure, but other implementations have also used more realistic representations [26]. Studies have found preferences for these briefer, visual scales when compared to more repetitive traditional EMA [26,38].

Mobile Sensing Approaches for Supporting Mental Health and Binge Eating

The ubiquity and sensing capabilities of smartphones make them attractive tools to passively collect multimodal sensor data 24/7. Compared to EMA, they are objective and less burdensome, have a higher temporal resolution, and provide rich data streams to infer aspects of users' social context and behaviour in naturalistic conditions [39,40]. Research has shown the potential of this data to monitor and support mental health conditions including depression [15,38,41–44], schizophrenia [13,45–48], bipolar disorder [14,49–53], stress [54–58] and anxiety [59–62]. Typically, behavioural features (metrics quantifying aspects of individuals' routines and activities) are computed from smartphone and wearable data and their role in tracking, classifying or predicting events of interest (eg, depressive states, hospital readmission) is explored via analytical methods. Clinical scales, medical records, or patient-reported outcomes are often used as a ground truth to validate the models built on top of behavioural features. However, to our knowledge only Juarascio et al. [16] have explored mobile sensing in the context of eating episodes associated with negative emotions. They monitored 21 people with clinically-significant emotional eating behaviours for four weeks using a wrist-worn device. Results showed that time- and frequency- domain features of heart rate variability can be used to classify 30-minute periods with and without emotional eating episodes, better than chance. Although they showcase the importance of wearable data in supporting binge eating monitoring, patient perspectives on mobile sensing and the role of smartphone phone data remained unexplored.

For eating disorders more broadly, research has shown that smartphone apps could increase patients' access to treatment [18], due to the anonymity they afford when considering the barriers people face to seek clinical help (eg, shame, fear of stigma)

[5]. Furthermore, in light of near-ubiquitous smartphone use in modern society, these devices are uniquely positioned to support access to resources, by promoting help-seeking and self-management behaviours [18,63]. Smartphones can enable personalised monitoring, which can aid in the identification of high-risk situations, derived from behavioural and situational context extracted from multimodal data. Alongside their capabilities for digital intervention provision, they offer a powerful platform for delivering support at optimal times [17,18]. Currently, there are a number of apps, designed primarily for people with disordered eating behaviours, which have been studied in the literature, including 'Recovery Record' and 'RiseUp' [18,64]. Both apps employ self-monitoring techniques and provide users with a set of coping strategies to try. In particular, 'Recovery Record' uses EMA to facilitate self-monitoring [65,66] and has some features which are similar to our DeMMI app (e.g., the ability to track mood and episodes of binge eating). However, it requires a significant amount of active tracking from the user (i.e. daily diaries, logging of meals and the feelings surrounding these), which can be a laborious task. Our study was interested more in how passive approaches to monitoring could be leveraged, allowing us to potentially automatically detect contextual or situational triggers for episodes of binge eating. This would ultimately remove some of the tracking burden from the user and provide indications for where digital interventions might be best positioned to help them. As a first step toward this goal, our paper explores the individual differences in smartphone behavioural features between days with and without binge eating episodes, framed around the experiences and needs of our users; as well as current and future challenges the mobile sensing community faces in our path to detect binge eating episodes and deliver digital interventions.

Methods

Our study was conducted in two stages. The first involved the development of the DeMMI app. As part of this, we conducted a consultation session to explore whether the types of sensor data we were proposing to capture were seen to be useful and appropriate, as well as gathering feedback on some specific app features relating to self-report. The second stage involved carrying out a 6-week period of data collection with 10 participants experiencing binge eating (logging both their mood and episodes of binge eating) and 10 comparison participants (logging only mood).

We first present the ethical considerations for this study. We then present the two stages of research separately, first describing the development of the DeMMI app, before moving on to discuss our fieldwork study methods and findings.

Ethical considerations

Ethical approvals for this work were obtained from York St John, UK, University Ethics Committee and adhered to the British Psychological Society ethical guidelines. The activities for the stage 1 consultation session were constructed collaboratively within the research team, made up of a clinical mental health professional and multiple highly experienced researchers, with expertise engaging people with a range of mental health issues in qualitative workshops and interviews.

The session was led by one of these experts and supported by an experienced post-graduate student working in the space of disordered eating and self-harm. Both facilitators were careful to create an open and non-judgemental space during the session. For stage 2, the research team created a safeguarding protocol prior to the commencement of remote participant recruitment. Participants were fully informed that their data was not being actively monitored and that the research team were not mental health professionals, however, we conducted weekly wellbeing checks via Whatsapp or SMS text messaging (depending on participant preference) during the study. These asked participants how they were managing with the study and offered an opportunity to reach out for support if required. Whilst we did not have any requests for support during the study, we were prepared to point participants to local services.

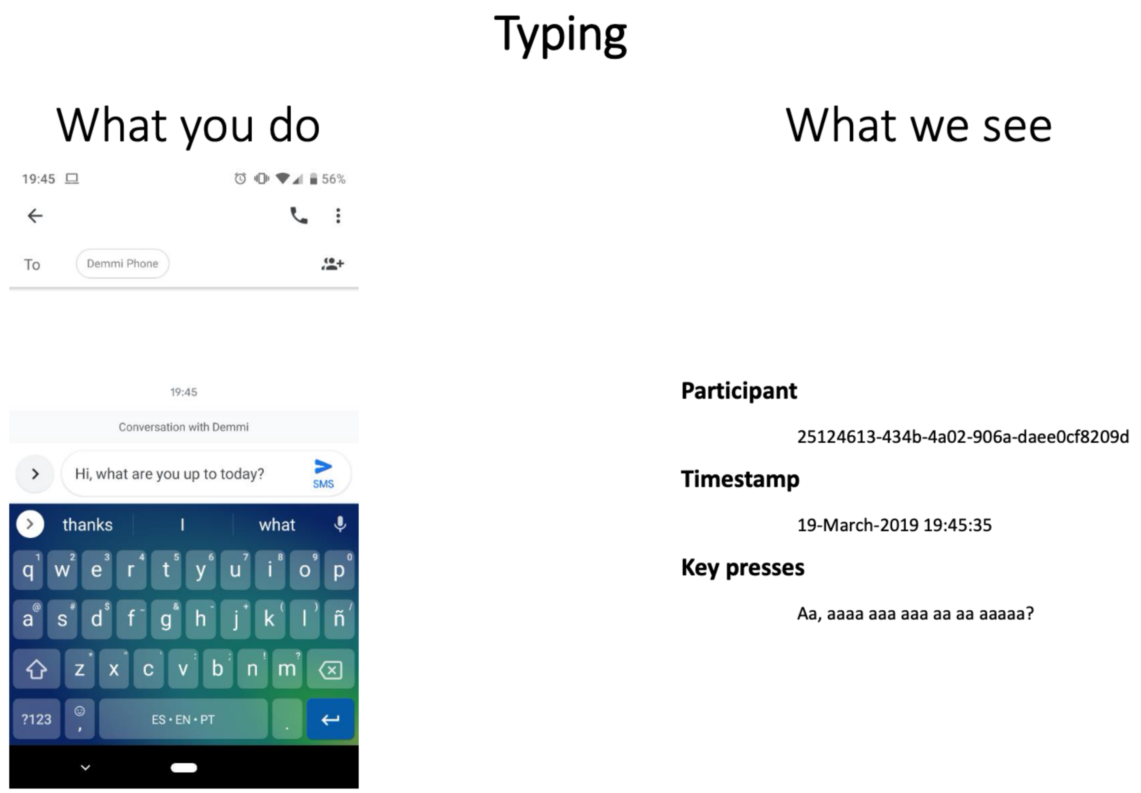
Phase 1: DeMMI App Development

Given the sensitivity of the data we wanted to collect, and an acknowledgement that reporting on disordered eating behaviours might, in itself, be considered as a trigger, we first conducted a consultation to gain an understanding of: 1) early perceptions toward the data we were intending to collect with the app, 2) gather ideas on the best ways to collect self-report data in a sensitive way, and 3) understand any specific opinions potential users might have around the rate of data capture, anonymity, and offers of support. We report the main insights drawn from this consultation to provide context for our design decisions, before moving on to describe the DeMMI app itself.

Consultation Activity

We engaged service users (n =2) from York, UK, in a consultation session to develop and iterate upon key design decisions related to the DeMMI app. One service user had lived experience of binge eating and the other had lived experience of self-harm (since a parallel project aims to explore the feasibility of DeMMI app use among individuals who self-harm). The consultation explored participants' perceptions of mobile data collection through both passive and active means. In terms of passive data capture, we explored participants' general perceptions of mobile-sensed data collection and also sought specific feedback on the intended data collection we were proposing within the study. Specifically how they would feel about us collecting potentially invasive sensor data. To facilitate this activity we used a bespoke "What You Do/What We See" resource (see Figure 1. The full resource can be viewed in Multimedia Appendix 1), which showed participants the output of every sensor. We discussed each sensor in turn and responded transparently to any questions. In terms of active data capture, we explored participants' general perceptions of using apps to log mood and behaviour, as well as their specific feedback on our proposed app logging mechanisms in the DeMMI app. There were two specific features we were interested in gaining feedback on. First, we were interested in understanding service users' perspectives of the use of the Positive and Negative Affect Scale (PANAS, which is widely used in EMA research with the general population [33–35]), to assess mood multiple times each day. Second, we were interested in the use of a one-click logging mechanism of binge-eating episodes based on the "oops" button developed by [67]. Data during the session was audio recorded to allow the team to listen back to the session, however as we only had two participants we did not thematically analyse the data. Instead, we took notes during the session and cross checked all key findings with the participants.

Figure 1. Example of one of our "What You Do/What We See" slides used to explain to participants what smartphone data we wanted to capture (ie., the participants' interaction) and exactly what we would see from this interaction. Slides were created to represent the following: app notifications, app use, typing, battery, calls, SMS, data sent or received, screen locks or unlocks, time zone, Wi-Fi nearby devices, Bluetooth nearby devices, ambient light, weather, location, ambient noise, and activity recognition data).



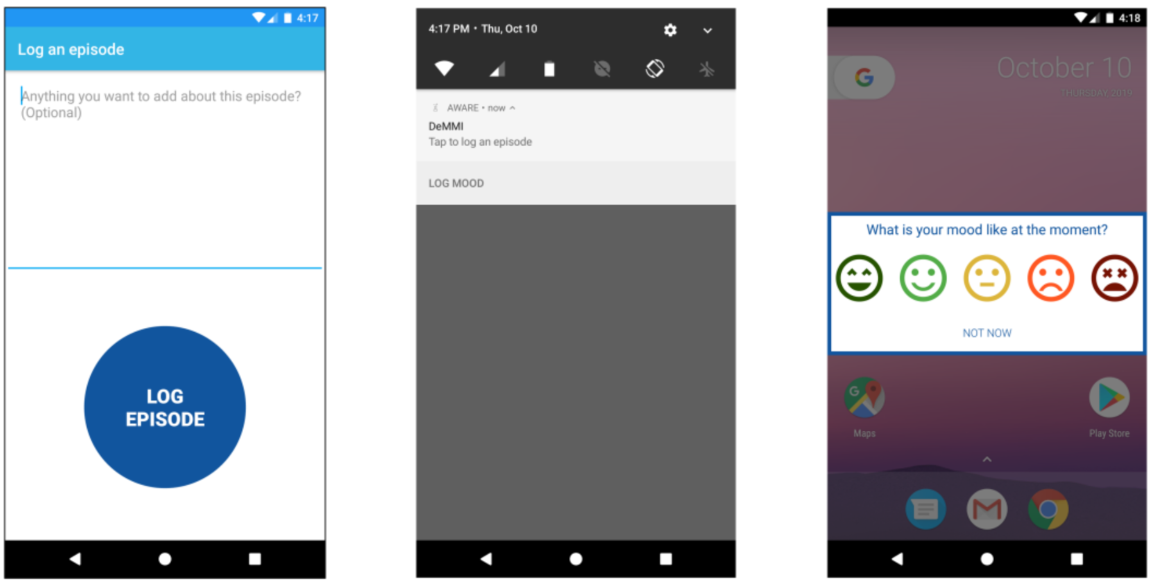
The key findings from the consultation may be summarised as follows:

1. Participants were generally positive about the use of mobile-sensing data for research and intervention purposes, identifying multiple beneficial uses including trigger identification, improved (i.e., more accurate and less effortful) behavioural logging, crisis management to prompt positive behaviours and support self-reflection. Any concerns they had regarding passively sensed mobile data (e.g., "that's creepy"), were overcome once data privacy had been reassured using the "What You Do/What We See" activity.
2. Discussions of the "oops" feature were mixed and led to modifications. Participants liked the idea of an easily identifiable, user friendly, and low effort button to quickly and efficiently record any incidences of problem behaviour, but felt that labelling this "Oops" was condescending, whereas "Log" was found to be more amenable. Participants also described how this method of behaviour logging could be improved by allowing the optional entry of descriptive text, either at the time of making the log or afterwards, so that they could add additional contextual information.
3. Finally, participants expressed concerns regarding the use of the PANAS, both in terms of the content of the scale and the frequency of proposed use in the study, e.g., *"If I had to do this three times a day, it would be a bit of a trigger for me…Just the word, ashamed, like if I had to go over that three times a day, I probably wouldn't do it, to be honest"*. A simple way to log mood (e.g. using smiley faces) was suggested as a preferable alternative.

DeMMI App Development

The DeMMI App is based on the AWARE [68] client for Android 7.0 or newer that we used to collect smartphone sensor data 24/7 from 14 sensors. The AWARE framework is an open-source mobile sensing platform used in context-aware mobile computing research. We collected data on accelerometer, app usage, notifications, battery, Bluetooth, calls, keyboard events, light, location, SMS, physical activity, screen power events, screen touch events and Wi-Fi data. Based on our consultation, we modified the client to allow participants to log binge eating episodes using a "Log episode" button placed below a textbox for open-ended feedback related to the episode- see the left screenshot of Figure 2; both were shown after tapping on the main body of a persistent notification labelled "DeMMI. Tap to log an episode"- see the middle screenshot of Figure 2. Participants would also be able to log their mood using a scale with emoji faces ranging in expression and colour to visually represent affective states on a 5-point scale ranging from very positive (represented by a very happy, smiling face) to very negative (represented by a very sad face)- see the right screenshot of Figure 2. This scale was automatically shown on screen every day at 9:00 AM and 9:00 PM and as soon as a participant tapped a face, the app logged the choice and hid the instrument. We provided a "Not Now" button so participants could ignore the prompt, and also allowed them to report their mood outside the scheduled times by tapping the bottom area of the persistent notification labelled as "LOG MOOD".

Figure 2. Screenshots of our button to self-report episodes (left), the persistent notification that allowed participants to open the episode and mood reporting screens (middle), and the mood reporting strings with five face emojis (right).



Phase 2: Fieldwork

The COVID-19 Context

This study was conducted with participants located in England, UK, between June 1st and August 14th 2020, as the COVID-19 lockdown restrictions were being gradually eased. Prior to this point, individuals had not been permitted to leave home except for limited purposes (e.g., shopping for essential items, exercise, medical care), and all non-essential shops, libraries, places of worship, and playgrounds were closed. During the time of our study, changing restrictions permitted that individuals from more than one household could meet outdoors in groups of six while maintaining 2m social distance (June 1st) and individuals living alone could form a 'support bubble' with one other household (June 13th). From July 4th, premises reopened with strict social distancing and hygiene measures in place and gatherings of up to 30 people were allowed both inside and outside private dwellings. Those facing disordered eating behaviours are thought to be particularly vulnerable during the pandemic. A study by Branley-Bell & Talbot [69] found that COVID was having a profound negative impact on people with Eating Disorders, while Schelgl et al [70] found 49% of patients reported a deterioration in Eating Disorder symptoms due to COVID-19 and 47% of binge eating patients reported an increase in binge eating symptoms.

Participants

Participants were recruited via several channels, including social media, the regional branch of a UK-based mental health charity (York Mind), and York St John University research participant recruitment forums. The study was advertised as a mobile sensing study for mental health, recruiting participants aged 18+ who currently reside in the UK and have an android phone as their primary device. Participants were recruited into two categories, 1) those with experiences of binge eating, defined as: "Eating an amount of food that you consider excessive, usually very quickly during a single session, eating until you feel uncomfortably full, eating when you're not hungry, eating alone, eating secretly, and feeling depressed, guilt, ashamed or disgusted after eating. This is often, but not always, accompanied by behaviours to counter the binge-eating (e.g. skipping meals, vomiting, over-exercising)"; and 2) those with no current mental health difficulties (comparisons). Participants with experiences of binge eating were not required to have a clinical diagnosis to take part in the study. Participants registered their interest by either directly emailing a designated member of the research team or completing an online form. Prospective participants were emailed the study information sheet and consent form and could ask any questions via email or phone call. All participation in the study was conducted remotely. A total of 20 participants took part in the study (10 with experiences of binge eating and 10 with no history of mental health issues). All participants were aged between 18-36 (mean age 25) and were almost exclusively female (with 1 male in each group).

Study methods

Once recruited into the study, participants were sent an onboarding pack via email. This included a link to the DeMMI app APK (an android package file format allowing the app to be downloaded via the link) and a set of step-by-step instructions for how to download, open, optimise battery life, and set permissions on the app (see multimedia appendix 2). Instructions on how to join the study with a unique identifier were also provided, as were instructions on how to log mood, log episodes and uninstall the app.

Participants were asked to run DeMMI for a total of 6 weeks, but were instructed that they could uninstall the app at any time, and that following removal all collected data would be deleted from their phone. All participants were asked to log their mood twice a day (at 9:00AM and 9:00PM) and the binge eating group were asked to log any episodes of binge eating using the "Log episode" button, with the option to provide free text information regarding the episode. To enhance engagement in the study and ensure that safeguarding protocols were being followed, participants were contacted once a week via text message to flag any potential problems. Following the study, participants were given the opportunity to take part in an optional interview, to discuss their experiences during the study and provide any feedback regarding how we might improve the app functionality in the future. Interviews were conducted via telephone or video call (depending on participants' preferences).

Data analysis

Quantitative

Smartphone behavioural feature analysis was conducted on the data collected from the binge eating group, with comparisons across days that an episode had been reported and days that an episode had not been reported. We used the Reproducible Analysis Pipeline for Data Streams (RAPIDS) [71,72] to pre-process, clean and extract behavioural features from the smartphone data we collected with the AWARE Framework. RAPIDS is a reproducible pipeline that allows the processing of mobile sensing data. According to our protocol, every participant had to be monitored for 42 days (i.e., 6 weeks) but in practice, their smartphones can run out of battery, our sensing app can crash, or it can have issues synchronising the data. Therefore, we expect some of these days to be missing all, most, or some of this data. We measured the quality of our smartphone data through the concept of valid sensed days; we labelled a sensed day as valid if we had 8 hours of data with at least 30 sensed minutes each. A sensed minute is a 60 second window with at least one row of data from any smartphone sensor.

Once data was processed, we used the Nonoverlap of All Pairs (NAP) index [73] to measure the probability that a behavioural feature value drawn at random from any episode day will exceed that of a feature value drawn at random from any non-episode day. NAP analysis provides an indication of effect size, offering directions for future work which might employ such sensor driven approaches in this context of binge eating.

Qualitative

All free-text episode logs that were collected during the study were collated and subjected to content analysis [74], to explore any recurrent themes of discussion across participants that might provide future directions for focus. Our analysis involved assigning codes to lines of the data and grouping these into themes. Interviews were all audio recorded and transcribed verbatim. Interview transcripts were thematically analysed, using a deductive approach which saw codes created at the paragraph and then sentence level [75].

Results

In this section, we describe and summarise the smartphone data, mood survey notifications, mood scores and binge eating episodes data we collected. In order to explore the relationship between smartphone data and binge eating episodes, we computed 12 behavioural features related to social interaction and physical activity across location, Bluetooth, physical activity, and screen sensor data and analysed the difference in their values between days with and without self-reported binge eating episodes. We share the code of our mobile app and analysis pipeline to enable the reproduction of our methods [76,77].

We present our analysis in 3 parts. 1) We first look at the entire cohort of data for all 20 participants (both the control group and the binge eating group) and provide an outline of the valid sensed days that we were able to collect (providing an indication of how long participants in both groups engaged in the study before deleting the app and how well the software functioned in relation to collecting the sensor data). Within this cohort level data we also provide a comparison of mood scores across all 20 participants to explore if there was any difference between mood reporting across the two groups. 2) In the second part, we look specifically at the binge eating data and the differences in sensor data collected on days where an episode was reported vs days where an episode was not reported. We first provide the content analysis of binge eating episodes and then discuss the NAP analysis that was conducted. 3) Finally, we report the interview data across both groups, which focused on the usability and acceptability of our approach and the app itself.

Part 1: Entire Cohort Data

Valid sensed days across all participants

Participants were asked to keep the app running on their phone for a total of 6 weeks (42 days). In relation to study engagement the two groups were relatively similar; the binge eating group had a total of 395 total days with the app, with 2 deleting the app before the end of the study (P02BE after 15 days; P06BE after 36 days). The comparison group had a total of 397 total days with the app with 1 participant deleting the app before the end of the study (P13C after 5 days). There was some difference in the valid sensed days however, with the binge eating group having only 72.4% of valid sensed days (n=286) compared to the comparison group who had 61.7% of valid sensed days (n=245). See table 1 for a full breakdown of data.

Table 1. Participants in the binge eating and comparison groups along with the number of days they were monitored and the number of valid sensed days

Participant	Group	Days with the app	Valid sensed days
P01BE	Binge eating	45	0
P02BE	Binge eating	15	12
P03BE	Binge eating	44	42
P04BE	Binge eating	43	31
P05BE	Binge eating	41	38
P06BE	Binge eating	36	6
P07BE	Binge eating	43	38
P08BE	Binge eating	43	39
P09BE	Binge eating	43	41

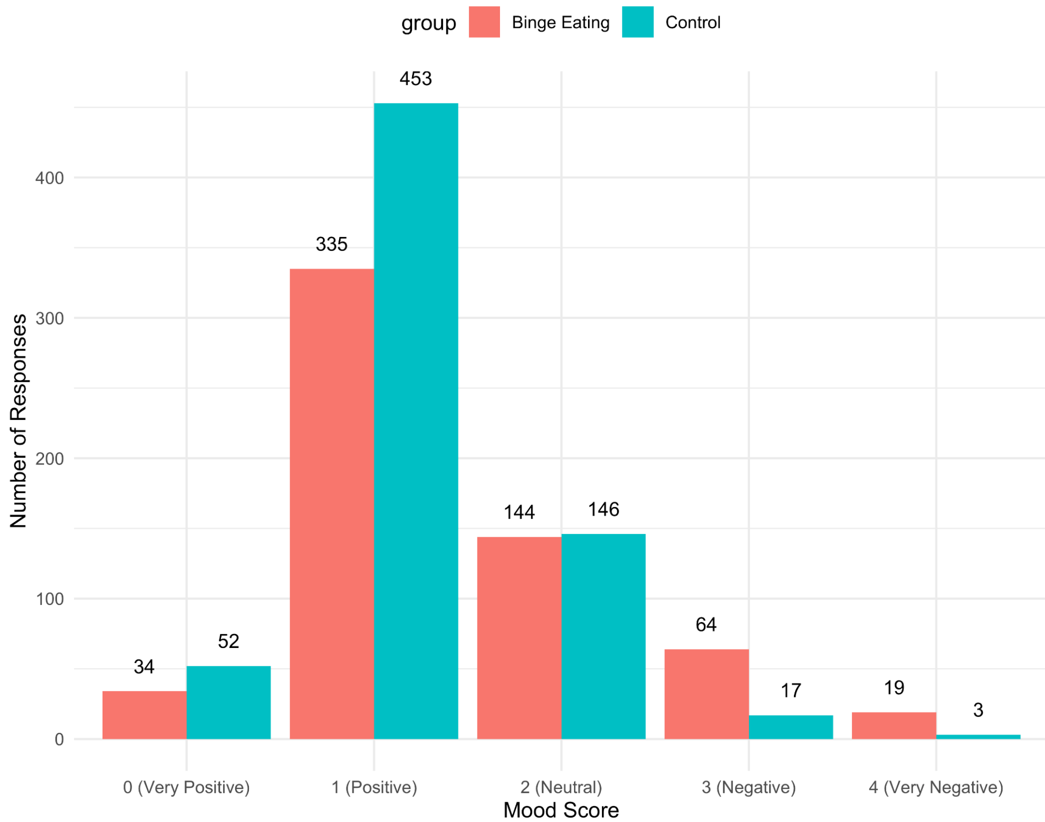
P10BE	Binge eating	43	40
P11C	Comparison	44	31
P12C	Comparison	44	35
P13C	Comparison	5	0
P14C	Comparison	43	40
P15C	Comparison	44	9
P16C	Comparison	44	20
P17C	Comparison	43	42
P18C	Comparison	43	2
P19C	Comparison	44	27
P20C	Comparison	43	39

Mood logs across all participants

Mood logs were collected twice daily; at 9am in the morning and 9pm in the evening. Overall the binge eating group logged their mood prompted by our app notifications a total of 548 times (of a possible 629) giving them a 87% log rate. The comparison group logged their mood a total of 638 (of a possible 706), giving them a higher rate of 90%.

However, as can be seen in Figure 3, the binge eating group had lower overall positive mood score, and were more likely to score their mood as negative (10.7%, n=64 compared to 2.5%, n=17 of the comparison group) or very negative (3.1%, n=19 compared to 0.4%, n=3 of the comparison group). The binge eating group were also more likely to report a lower mood log in the mornings (11.4%, n=34) and evenings (13.1%, n=39) reporting negative or very negative scores compared to comparisons (1.82%, n=6 & 4.13%, n=14).

Figure 3: Bar chart comparing the mood score categories from 0 (very positive) to 4 (very negative) for the comparison and binge eating groups



We computed the rate between the number of mood surveys that were displayed on screen and the number of surveys that should have been triggered (2 per sensed day). Our participants had a mean trigger rate of 92.9% (SD=10, Min=59.4, Max=100) which shows our surveys were delivered reliably. We also calculated the rate between the number of surveys that were answered and the surveys displayed on screen. We got an average answer rate of 88.4 (SD=6.71, Min=67.5, Max=96.2) suggesting that our mood survey instrument was easy to use which is further confirmed by the fact that the number of ignored surveys was on average 1.3 per week through the study's six weeks. We also gave participants the opportunity to report their mood outside the scheduled time. This functionality was rarely used, with a total of 95 self-initiated mood reports and an average across all participants of 5 responses (SD=5.9, Min=0, Max=18). However, we highlight that 26.3% (25) of these surveys were used to correct a previous self-reported mood score (defined as any new report made within 2 minutes of the previous one).

Part 2: Binge eating Data

Episode logs

We had a total of 98 episodes reported across our 10 participants in the binge eating group (see table 3). Most logs occurred in the evening between the hours of 6PM-12AM (n=41, 42%) or in the afternoon between 12PM and 6PM (n=35, 36%), with a smaller number occurring in the morning between 6AM and 12PM (n=14, 14%) and a small amount happening late at night between 12AM and 6AM (n=8, 8%).

Table 3: total number of episode logs and episodes logged with additional text for all binge eating participants

Participant	Total number of episode logs	Episodes logged with additional text
P01BE	6	0
P02BE	3	0
P03BE	9	3
P04BE	6	2
P05BE	8	7
P06BE	23	23
P07BE	19	18
P08BE	16	2
P09BE	6	5
P10BE	2	2
Totals	98	62

Of the 98 episode logs, n=36 (37%) were recorded without any additional text. We conducted a content analysis on the remaining n=62 self-reported episodes with additional text, to explore any related themes of reporting that cut across participants. Please note that some episodes reported multiple themes. There were 6 overarching themes identified from our content analysis: 1) Perceived lack of control; 2) Emotions; 3) Invasive negative thoughts; 4) Disordered eating behaviour explicitly reported; 5) Situational influence; 6) Behaviour avoided.

We had n=36 explicit reports of binge eating behavior across participants: *"I was watching movies and binged snacks even though i'm not hungry"* (P05BE); *"Made coffee muffins when partner was out. Ate 6 in a row."* (P06BE); and *"massively over-snacked after not eating much lunch at 12pm"* (P08BE). There were n=6 reports which displayed a perceived lack of control: *"tried really hard to not binge, but couldn't stop myself"* (P03BE); and *"I'm on my own and I can't sleep, and I can't stop snacking"* (P05BE). Invasive negative thoughts were also explicitly reported n=12 times: *“Self-critical thoughts. Strong emotions”* (P04BE)*; “feel like failure.”* (P06BE); and *“I'm having really bad intrusive thoughts :(“* (P09BE). Moreover, participants reported on their feelings and emotions after a binge eating episode: *“feeling quite ashamed now”* (P03BE); *“feel useless and angry”* and *“Feel disappointed and dumb”* (P06BE). Participants also reported n=15 situational influences, which led them to disordered eating behaviors: *“Had to do a job interview today but didn't Go to plan”* (P03BE); “*jeans I had ordered arrived today. They didn’t fit and all I wanted to do was sit and cry and eat but I couldn't because I wasn't alone”* (P05BE); and *“every time I see an ad for that massive popcorn chicken from KFC I wanna binge eat the entire thing”* (P06BE). Finally, there were n=6 remaining reports that related to disordered eating behaviors that were avoided, despite the thought processes around binge eating being there, due to factors such as not being alone: *“want to make myself sick in the bathrooms but it's empty and I'm scared my partner will hear.”* (P06BE); or due to the acknowledgement of the negative effects of the disordered eating behavior: *“didn’t realise how much I was eating till I finished[...]but I am trying to stop purging as it is starting to have negative physical effects”* (P03BE). In the case of P06BE, there were several reporting incidents that were framed positively (n=5): e.g. *“Feel good didn't have cake etc at cafe as I would 100% scoff and eat too much and feel ill like i usually do”* (P06BE) which may have been perceived as restraint in relation to avoidance of a binge eating episode, however it is worth noting that deliberate restriction of food can also be problematic and a prelude to binging.

Smartphone data on episode vs non episode days

P01BE only had 1/45 valid days, P02BE only had 14/15 valid days, and P06BE only had 6/36 valid days as such, they were excluded from our behavioural feature analysis. This left us with the remaining 7 participants. We explored the difference in magnitude between days when our participants did and did not report an episode in relation to 2 self-reported mood scores (those logged in the morning between 6am and 12pm and those logged in the evening between 6pm and 12am) as well as 12 behavioural smartphone features extracted in daily segments that span midnight to midnight. The smartphone features we included were: geographical location variance, total distance travelled, radius of gyration, time at home, stationary (to moving) ratio, total stationary time, number of distinct Bluetooth devices sensed around the phone, number of screen unlocks, total screen time, time of first screen use, total number of mobile applications used, and the entropy of all mobile applications used (a wider variety of apps produce a higher entropy value). The sensors that provide these features, a layman’s description and implementation of these features can be found at RAPIDS documentation [71].

We framed our problem as seven n-of-1, or single-case, experiments with alternating AB phases. An A phase is formed by consecutive days without binge eating episodes while a B phase is formed by consecutive days with binge eating episodes. Overall, 81.8% (36/44) of B phases across all participants were 1 day long (mean 1.3 days range 1-4). A phases across all participants (n=48) were on average 5 days long (range 1-28). Table 4 displays a breakdown of average phase length and range for each participant.

Table 4. Average length of A (no episode) and B (episode) phases

Participant	# A Phases	A phases average length (range)	# B Phases	B phases average length (range)
P03BE	7	5.1 (1-10)	6	1.3 (1-2)
P04BE	6	6.3 (1-21)	5	1 (1)
P05BE	6	5.5 (1-13)	6	1.3 (1-3)
P07BE	9	3.1 (1-9)	9	1.7 (1-4)
P08BE	11	2.6 (1-7)	11	1.3 (1-3)
P09BE	6	6.3 (1-14)	5	1 (1)
P10BE	3	13.7 (1-28)	2	1 (1)

We used the Nonoverlap of All Pairs (NAP) index [73] to measure the probability that a behavioural feature value drawn at random from any phase B will exceed that of a feature value drawn at random from any phase A. We used the R implementation provided by Pustejovsky et al. [78] concatenating all (1-N)A phases into a single A phase and all (1-M)B phases into a single B phase, and setting a confidence threshold of 0.95 and an expected direction of improvement given by the sign of the standardised mean difference of A and B values ($[\text{mean}_B - \text{mean}_A] / \text{sd}_{[A,B]}$).

We share in Table 5 the behavioural features of each participant that had a NAP medium effect (an index between 0.66 and 0.92) or a NAP strong effect (an index between 0.93 and 1) [73]. Only the strong effects belonging to P10BE were statistically significant after adjusting for multiple tests within participants using the Benjamini and Hochberg method [79]. P04BE, P05BE, P07BE, and P10BE self-reported to feel worse on days when they binge eat, P04BE having morning mood at a medium effect, P05BE and P07BE having evening mood at a medium effect, and P10BE evening mood at a strong effect. P05BE, P08BE and P09BE had screen features at a medium effect (P08BE and P09BE used or unlocked their phone less in days when they binge eat). P03BE had app entropy at a medium effect (he or she used a wider variety of apps on days when they binge eat). P04BE and P10BE had time at home at a medium effect (they spent more time at home when they binge eat). P04BE, P05BE, and P10BE had total distance at a medium and strong effect (P04BE travelled more and the other two less in days when they binge eat). P09BE and P10BE had a stationary time and ratio at a medium effect (they moved around less when they binge eat). Finally, P10BE had location variance and radius of gyration at a strong effect (he or she travelled around less when they binge eat) and unique bluetooth devices at a medium effect (he or she was arguably around fewer people or public places when they binge eat).

Table 5. Smartphone features that showed a medium or high NAP effect between phases A and B of participants that binge eat

	Feature	Std mean diff	NAP (SE)	NAP CI	NAP Effect	Adj p-value
P03BE	App entropy	0.67	0.69 (0.1)	(0.46, 0.84)	medium	0.451
P04BE	Time at home	0.64	0.73 (0.9)	(0.46, 0.89)	medium	0.450
P04BE	Mood morning	0.62	0.68 (0.1)	(0.41, 0.86)	medium	0.450
P04BE	Total distance	0.38	0.67 (0.16)	(0.40, 0.85)	medium	0.450
P05BE	Total distance	-0.58	0.73 (0.11)	(0.51, 0.87)	medium	0.285
P05BE	Mood evening	0.79	0.69 (0.13)	(0.47,0.85)	medium	0.285
P05BE	Screen unlocks	0.36	0.67 (0.08)	(0.45,0.83)	medium	0.334
P07BE	Mood evening	0.70	0.68 (0.08)	(0.50, 0.81)	medium	0.246
P08BE	Screen unlocks	-0.63	0.69 (0.08)	(0.51,0.83)	medium	0.292
P09BE	Stationary time	0.79	0.68 (0.16)	(0.41, 0.86)	medium	0.437
P09BE	Screen time	-0.81	0.67 (0.16)	(0.41, 0.85)	medium	0.437
P10BE	Mood evening	1.67	0.99 (0.02)	(0.58,1.00)	strong	0.046
P10BE	Radius of gyration	-0.46	0.98 (0.02)	(0.57,1.00)	strong	0.046
P10BE	Location variance	-0.32	0.98 (0.02)	(0.57, 1.0)	strong	0.046
P10BE	Total distance	-0.68	0.98 (0.02)	(0.57, 1.0)	strong	0.046
P10BE	Stationary ratio	1.47	0.91 (0.09)	(0.5, 0.99)	medium	0.074
P10BE	Time at home	1.01	0.88 (0.06)	(0.46, 0.98)	medium	0.091
P10BE	Bluetooth devices	-0.88	0.83 (0.07)	(0.42, 0.97)	medium	0.125

Part 3: Post-study Interviews

Post study interviews were optional. Among the 10 binge eating participants, 3 consented to the post-study interview; among the 10 comparison participants, 5 provided consent, giving a total of 8 interviews for analysis. Interviews lasted between 14 and 42 minutes each, all were conducted via telephone and were audio transcribed for later deductive thematic analysis. A total of 45 codes were initially created and were then further grouped into themes. There were four overarching themes identified: 1) Positive and Negative Impact of Lockdown; 2) Phone Habits; 3) Mood and Episode Logging; and 4) Usability of DeMMI App.

Positive and Negative Impact of Lockdown

In the interview, participants were asked how their mood might have changed during lockdown. The majority of participants discussed their mood negatively: *"I was doing a lot worse when lockdown got really bad."* (P09BE), with both groups reporting fluctuations in their mood during lockdown: *"it was very much up and down…"* (P13C); and possibly triggering episodes for those in the binge eating group: *"I definitely think it's been a lot more up and down…as soon as something tiny goes wrong…I would have like a full-on breakdown. Then I'll comfort eat, have a binge or start like picking at myself"* (P07BE). Several recurring themes were noticed in the interviews, which related to negative impacts of lockdown on one's mood. Some participants highlighted feelings of hopelessness: *"It can feel quite hopeless as we don't know when it will change"* (P10BE); and that the uncertainty of the situation was causing stress and anxiety: *"I think like everyone, it's had a negative impact…there was that sort of novelty factor and we weren't quite sure how long it was gonna last."* (P11C); and "*it was definitely a bit more scary, and because we didn't know how long it was gonna go on for. A bit unsettling…that was a bit stressful"* (P15C).

However, there were also some positives reported by both groups. Two participants mentioned how lockdown had encouraged them to have better time management, since the time usually spent getting ready and commuting to work could be used to spend more time with people around them and activities they enjoyed, such as exercising and volunteering: *"there's been a lot of benefits of lockdown… I managed to get into a nice little rhythm once my routine kinda reset"* (P15C); *"it's [lockdown] given everyone a focus which in some ways has helped my social anxiety, I've been volunteering which I wasn't before"* (P10BE). Furthermore, while some participants highlighted that being able to exercise *"makes [them] feel good and it kinda like refreshes and resets [them]"* (P1C), for members of the binge eating group *"not being able to go and do exercise as much, just feeling very tired all day"* (P07BE) was often seen to be a cause of stress which could trigger a vicious cycle of binge eating episodes and dietary restrictions. Despite this, the same binge eating participant reported how having regular face-to-face social interactions during and post-lockdown led to a positive impact on their mood: *"if I go into placement. I work well, it's a good day, but if I'm at home it's not a good day"* (P07BE), but with lockdown limiting face-to-face contact, participants in this group could be at an even higher risk of triggering a binge eating episode.

Phone Habits

Long periods of staying at home during a pandemic, with limited knowledge of the virus, was seen to take a toll on participants' mental health and with restrictions on socialising, participants discussed beginning to form new habits, especially with their phone and social media usage. During lockdown, most of the participants reported a noticeable increase in their phone usage: *"I've been on my phone a lot since lockdown started."* (P09BE); *"I definitely use my phone a lot more than I used to… I wouldn't have done as much if I didn't have as much free time as I do now."* (P15C); and *"I would say that I was probably spending a bit more screen time in lockdown…Not having that structure."* (P14C). One participant clarified that this was not due to taking part in the study: *"I think it's more lockdown that's increased my phone behaviours not really the app."* (P07BE). The major causes for this change in phone use behaviours noted by many participants were procrastination: *"A lot more procrastination, just not really using it for anything useful."* (P07BE) and *"I'm into the habit when I'm not doing anything, I'm much more procrastinating on my phone than I used to be."* (P09BE); and boredom: "*I was using that a lot more, and maybe even out of boredom."* (P13C); and *"maybe out of boredom a little bit and maybe just out of habit as well."* (P14C), particularly since the lockdown period was during the summer break for university students and/or some participants were forced to take a break from work.

A few participants also discussed the negative impact of social media and phone use on one's mental health: *"the increase in my phone usage probably contributed to the decline in my mental health."* (P10BE); and *"when I'm on social media and I see negative things. So, when I get news alerts it's always negative… I think Twitter was the worst for me…when everything was going on with coronavirus and like the Black Lives Matter movements…everything really made me feel down and depressed."* (P17C). Moreover, three of the comparison group participants discussed the positive impact of using the DeMMI app: "*it made me think about my use of social media or being on my phone and if I thought that correlated with my mood…I was thinking about how work affected my mood, rather than my usage of my phone."* (P14C); *"I did think about phone usage as well because when I was thinking about what the app is sort of looking at, I would reflect and think that actually I probably use my phone a lot more than I thought I did."* (P13C); and *"I noticed how often I was using my phone, and I was quite conscious of it in the first couple of days, I did notice patterns in my mood too."* (P15C). It seemingly led some participants to reflect how their phone usage would correlate with their mood.

Mood and Episode Logging

The binge eating participants were asked to expand on what they felt triggered them to experience an episode, and how they classified an episode during the study (please note again that the comparison group were only instructed to log their mood). For one participant, the impact of social relationships on the occurrence of episodes was noted: *"in terms of the binge eating…Cause, my boyfriend at the time…we had some sort of issues"* (P07BE); and the passing of a close family member had impacted on their mood: *"probably about 6 weeks ago there was a big trigger… my uncle had been ill for a while… And then he died about 6 weeks ago"* (P07BE). Other binge eating participants noted that *"intrusive thoughts"* (P09BE) and *"the increase in my phone usage probably contributed to the decline in my mental health"* (P10BE), which might trigger episodes of binge eating for them. In the interview, participants were also asked whether they have noticed any patterns in their behaviour. Two of the binge eating participants also noticed that their mood had improved with the return of social interactions and having more occupied time: *"So now it's a lot less time for overthinking and worrying about things, and more time to actually just be doing stuff."* (P07BE). All the comparison group participants reported no concerns with mood logging. On the other hand, one of the binge eating participants found it invasive at times due to logging in a public environment: *"Sometimes it was annoying because I wanted to look at my phone for something else. Sometimes if I was with someone, I didn't want them to see what I was clicking so I didn't do it."* (P10BE). However, another binge eating participant noted that the app provided them with a level of accountability over binge eating episodes *"having to log it and recording my episodes has actually been really helpful as like a deterrent, especially for my binge eating. Because it was giving me some accountability for doing it and if I do binge eat then I have to record it. I actually think that it made my like threshold for having an episode a little bit higher."* (P07BE).

Usability of DeMMI App

When discussing the usability of the DeMMI app, participants provided mostly positive feedback for the app. Many participants commented that it was "*easy to use*" (P13C, P14C, P15C, P17C, P10BE). Some of the participants elaborated on the user-friendly aspects of the app: *"I liked [how] you could just click on the notification bar and then just tap to add episode"* (P07BE); and *"I liked how it popped up with the pop up telling you to log your mood now. Because otherwise 100% I would have forgotten."* (P09BE). The main issue reported by participants was related to the app crashing. Two of the participants commented that "*I had a lot of issues with the app force closing and not knowing if it was still running, I don't think in the end I could enable all the features of it."* (P07BE); and *"The only thing that did happen, was occasionally when I was browsing the web, so not really when I was using any apps, it would occasionally say the app wasn't responding."* (P15C).

Participants provided several specific examples of feedback that could be used for improving the app in the future. Four participants proposed an option for additional mood logs, one participant suggested, *"Maybe one around lunch time and one around evening maybe?"* (P09BE); another elaborated with a similar suggestion saying, *"Because there was lots of days where in the morning and evening I would put a neutral or a happy face, then in the middle I would dip to close to having an episode but not quite have an episode."* (P07BE). P10BE commented that they would like the ability to adjust the time of the mood logs, saying *"I am often only just up by 9am or still asleep, so checking in later for a mood score would have been better…sometimes I needed to urgently look at something on my phone, so I clicked off it quickly without thinking and missed logging it."*

All the binge eating participants could see the benefits for a future app version to identify or predict a potential episode: *"even it's just recognising you're about to have an episode… I think that might in itself even be a bit useful. Like giving the accountability before and after."* (P07BE), as well as the addition of mindfulness and relaxation exercises or safe practices: *"if I pre-put in some like songs or something that I like and then it would identify it and then prompt you to play the song, rather* *than needing to go and think about oh I'm going to put on some relaxing music. Or maybe some exercise which uses your mind a bit to distract you."* (P07BE); and *"Maybe if you could tailor them to what you find useful. I wouldn't like an app to tell me what to do. Maybe if you'd set up a reminder to do mind*fulness or something." (P10BE); or positive affirmations: *"like positive affirmations…Like, a flowchart on what to do."* (P09BE).

Lastly, participants were asked about privacy concerns regarding the automated data collection by the app. Most participants voiced no privacy concerns: *"I kinda trusted the study"* (P13C). Moreover, two participants explained, *"but if this was just another app on the app store for sure I would be having privacy concerns."* (P09BE); and *"no [concerns] because I was told you could see what apps and stuff were open, but you couldn't see any messages. I was worried initially when I was signing up to the study, but after I read everything and emailed you it made me feel a bit better about it."* (P17C). Further open questions were asked surrounding the clarity of the information sheet that was distributed to participants. A majority of the participants were satisfied with the level of information that was provided for them to understand the purpose of the study and how their data would be collected and handled: *"I didn't have any concerns or anything about that…I think it was very clearly set out, I don't think anything stood out particularly."* (P14C); and *"No [concerns], I was really happy with it…in terms of privacy there was nothing that really stood out in the information sheet that I was, ooh I'm not sure about that or I don't understand that…it was all pretty straight forward."* (P14C). Several participants further highlighted the importance of explicitly discussing data privacy and security within the study: "*the information sheet where they were pointing out the privacy settings, I think that stood out the most to me. Because that was the information that I wanted to know the most before I started the study."* (P17C); and "*I did at first, just in terms of there's so much data that you can really get passively from someone's phone…But I knew what was being done with the data, so I wasn't too concerned about it."* (P15C).

Discussion

Our preliminary quantitative results suggest that every participant had various smartphone features that are meaningfully different between days with and without binge eating episodes. This in turn could encourage researchers to investigate fully data-driven approaches to find "hidden" links between smartphone behavioural features and these episodes either via interpretable or black box predictive approaches. However, our qualitative work paints a more nuanced image of the research needed to deliver effective, safe, and ethical digital interventions.

Perceived Episode Indicators

In the feedback attached to the self-reported episodes, participants described a variety of affective states, comorbid mental health disorders, social interactions and daily life experiences that either preceded or happened during their binge eating episode. These findings correspond with the existing research literature surrounding binge eating episodes, which has also identified these factors as potential antecedents of binge eating episodes [80]. We refer to these as indicators, due to their predictive potential, and are keen to emphasise that indicators are not necessarily causal contributors to binge eating episodes. It may be that they are provoked by the latent trigger itself, for example, calling a relative for support may co-occur with binge eating but might not necessarily cause the binge eating episode. Having in mind that the basic premise behind a digital intervention based on mobile data is to find behaviours that can be measured in order to deliver a treatment before, during, or after a binge eating episode, it is crucial to understand and catalogue these indicators independently of their causal role.

The idiosyncratic nature of indicators identified in this study warrants further examination in the form of a longitudinal observational study, which can help clarify the relevance, scalability and focus of quantifying indicators using mobile data (in our case screen unlock time, app entropy, time at home, stationary to location ratio, total distance, stationary time, screen unlocks, evening mood score and time of first screen unlock). For example, a scenario where most of these indicators are related to affective states or mental health disorders like depression, would support the idea of leveraging previous works on general or personalised mobile monitoring of these constructs [81]. Under a different scenario, if most of these indicators are related to situational influences like social interactions with certain people, work activities or leisure activities, then it is likely that models to detect these events will have to be highly personalised given the differences in people's routines. For example, researchers might have to monitor digital communications between participants and specific relatives or friends or adapt to people's work, school or leisure settings (monitoring sleep patterns, drinking patterns, physical activity, etc.). In practice, there might not be a clear-cut line between affect, mental health and situational influences as the latter is likely to affect the former, and a participant could report episodes around both types of indicators, however, some might be easier to quantify using smartphone or wearable data (see Computation amenability of indicators below).

It is also worth noting that the idiosyncratic nature of the indicators we have identified might reflect the idiosyncratic nature of binge eating more broadly. In this study, we defined binge eating in inclusive terms, did not state that this must form part of a specific diagnosis (eg, bulimia nervosa, binge eating disorder), did not exclude binge eating which is comorbid with another mental health condition (eg, anxiety) and did not specify whether or not the binge eating must occur in the presence or absence of purging. Past EMA research suggests that different factors may be more or less important to different clusters of participants. For example, a review by Dingemans et al [82] found differences in the affective dynamics associated with binge eating with purging compared to binge eating without purging. Thus, the previously described work on binge eating indicators would benefit from collecting this information in order to understand whether there are commonalities between certain participant clusters.

Frequency and Lifespan of Episode Indicators

We need to understand how often and for how long these indicators happen such that researchers focus their efforts on the most common ones for a participant or a cluster of participants. It is possible that the time a person is exposed to or experiences an indicator will vary and that their relevance will fluctuate over time. The former means that the timescale at which mobile data is analysed will depend on the indicator (eg, should we look at anxiety levels days or hours in the past?) and the later implies that models will have to adapt over time to changes in people's routines and personal circumstances (eg, if being alone used to trigger binge eating episodes, what happens when a young adult moves to live on their own after sharing a house with other people during university?). Further compacting this issue, research examining binge eating episodes has suggested that these may not necessarily be discrete events. More specifically [83] discuss how binge eating may be best thought of in terms of binge eating days rather than binge eating episodes. This is especially important where binge eating occurs without purge behaviour since the purge behaviour is often considered a clear indicator that an episode has finished [84].

Severity and Impact of the Episodes Around Certain Indicators

Paying particular attention to the situational influence indicators (eg, being alone or having an argument with a loved one), it is unclear a) how often they are a proxy to binge eating episodes (and their likelihood of triggering a false positive intervention); b) whether or not the severity of the episodes they pinpoint is similar as measured by objective means (such as the amount of consumed food, its nutritional value or the duration of the binge) or subjective means (such as the extent to which they experience lack of control); and c) whether or not the psychological and physical impact of such episodes is similar across episodes (e.g. people might not always engage in purging behaviours after an episode). Researchers and patients might prefer to investigate, quantify and monitor the indicators that pinpoint episodes with the most negative effect. Our lightweight approach to episode logging was well received by participants, who provided additional text in 63% of cases. Leveraging Natural Language Processing approaches to gain a better understanding of specific indicator types and their severity of impact on the person is a promising direction for future work.

Computation Amenability of Indicators

We expect to be able to quantify and detect each indicator to a different degree using smartphone or wearable data. In our study, we computed smartphone features that measured constructs we considered to be roughly related to the indicators reported by our participants. However, once researchers have a better idea of the breadth and depth of indicators in a population, they can decide what behavioural features might be more relevant and need to be extracted. For example, it might be very difficult to measure or anticipate the effect that daily activities like shopping will have on body image (P03BE) but if it turns out that a considerable number of binge eating episodes happen around body image issues, then we could focus on measuring the affective state induced by them. Alternatively, research has indicated that some individuals may engage with certain apps (eg, calorie-tracking app) in different ways around binge eating episodes [1]. Similarly, certain episodes could happen around indicators related to physical activity, sleep disorders, or communication patterns that previous research has had positive results quantifying using smartphones and wearables [47,59,85–87], which are a more direct measurement of the observed phenomena compared to affect and psychological constructs.

Intervention Candidates for Episode Indicators

Once a relevant indicator is identified and can be reliably measured, the next question is how to intervene to try to prevent the binge eating episode from occurring and/or provide support for its duration. There is a pressing need for effective binge eating interventions: currently available treatments are only effective for up to 50% of individuals with binge eating disorder and 30% of individuals with bulimia nervosa [88,89]. Of these treatments, the most commonly prescribed is Cognitive Behavioral Therapy (CBT), which involves restructuring an individuals' thoughts, feelings and behaviours in order to support more productive outcomes and reduce binge eating occurrence [90]. While CBT is typically delivered in therapeutic settings, research has begun to consider its potential when delivered remotely through smartphones (akin to a self-help tool) [17]. Initial research in this space suggests that smartphone-enhanced CBT can be as effective as therapist-led CBT, and maybe even more effective in attaining some outcomes, such as meal adherence [22]. Such smartphone-based approaches may also be useful at the subclinical level, where individuals binge eat and experience significant distress without diagnosis.

There are several types of intervention that may be appropriate for in-situ delivery in response to the mobile sensed occurrence of an indicator. These interventions can be loosely categorised as 1) prompting and 2) self management; both of which may be compatible with a CBT framework. Prompting interventions refers to those aimed at nudging participants to change their behaviour, prompting self-reflection that helps put things in perspective, or even makes them consciously aware that an indicator that is typically associated with their binge eating behaviour has been detected. Self management refers to interventions that support participants by providing access to online tools that foster positive mental or physical health. These online tools could take many forms including the automatic recommendation of activities aimed at de-escalating the situation (such as distraction activities available in apps like 'Calm Harm' or 'Recovery Record'), support messages that are meaningful for the individual, or open communication channels to family, friends or health care providers that the participant agrees to. Indeed, our participants provided similar recommendations for activities that might support them in future app iterations during the post-study interviews. Our empirical programme suggests that such interventions would need to be personalised to the participant, consistent with previous work on the design of smartphone-based interventions to support mental health [91]. This personalisation could be done with the support of a therapist (ie, a therapist mediated intervention), or be self-led (ie, a self-help approach).

Risks, Cost and Effectiveness of an Intervention

Researchers need to systematically consider the cost-effectiveness of prompting, self management and other kinds of interventions. This type of analysis has been carried out for HIV [92], physical activity [93], smoking [94], alcohol consumption [95] and CBT guided self-help interventions for binge eating [96]and should take into account the time and expertise that these kinds of digital interventions would demand from participants and their health care providers [17]. In addition, researchers need to be aware of the risks of delivering an intervention when it is not needed, failing to deliver an intervention when previous deliveries have been successful and participants rely on them, delivering an intervention aimed to disrupt a particular indicator that in turn puts the participant at risk of engaging in situations that could still trigger an episode (e.g., suggesting someone to avoid texting a relative without knowing that this could make them anxious and in turn provoke an episode) and the long term side effects of following certain interventions (e.g., spending less time outside). Designing future solutions which actively consider responsible innovation and the possible negative consequences of certain app features will help us to avoid any unintended consequences [1].

In the end, it is fair to assume that binge eating monitoring and digital interventions will need to account for the frequency, lifespan and computational amenability of the episodes' indicators, and find a trade-off between the severity and impact of such episodes, and the goal, cost, risk and effectiveness of candidate interventions. This sensor-informed context could support people and their therapists to identify triggers of their binge eating episodes or simply augment the non-digital strategies they already use. As with other forms of retrospective self-reporting [97–99], the consequences of showing participants historical contextual information, the validity of these reports, and cognitive biases that might come into play should be studied and to our knowledge are an open problem in binge eating research.

Usability and Acceptance

While we wanted to know how to improve the functionality of the DeMMI app in our interviews, we did not set out to follow a usability engineering method of evaluating the usability. Usability discussions were organically gathered from the interviews and were identified as one of the themes from our qualitative analysis. At this early stage of our research, where we were simply collecting data as unobtrusively as possible, we wanted participants to have minimal tasks to complete when interacting with the app; they were only expected to log episodes and/or mood. Poor system usability is highlighted as one of the factors affecting patient acceptance of health technologies [100]. As such, we envision the need for conducting usability evaluations such as heuristic evaluation and end-user testing, when the DeMMI app is further developed to include intervention based functionality. End-user testing examines how users carry out certain tasks or follow processes and is mainly focused on users' experience within the system [101]. Heuristics evaluation is carried out by usability experts and is concerned with the assessment of the system against a set of heuristics guidelines [102]. This type of usability engineering method would be an essential step when end users are expected to navigate through the app and engage with intervention based content aiming to support them in managing their disordered eating behaviours . Future work should adapt usability evaluation approaches, taken for example by [103] and [101], which are recommended for digital health solutions.

Reflections on Approaches to Participant Engagement

Participant insights played a pivotal role in uncovering the links between mobile data and binge eating occurrence. That said, participant engagement was a significant challenge that we experienced throughout our study. Our most successful recruitment medium was an Instagram campaign that 177 people responded to via email. However, only 15 people consented to participate after receiving the study information and out of those, only 10 installed our app. This low recruitment rate could be due to concerns surrounding the sharing of mobile data that has the potential to expose web browsing and communication habits. However, as was noted in both our consultation and post-study interviews, and echoing findings from [91], participants were generally satisfied with the information we had provided them relating to how we would capture and use their sensor data through our "What You Do/What We See" resource, which we have made available for reuse (multimedia appendix 1). This clarity and transparency regarding what could be considered invasive data capture was enough to alleviate participants' initial concerns and we greatly suggest that future researchers use a similar approach in their own research to increase participants' literacy surrounding mobile sensed data. Furthermore, only 3 of our 10 binge eating participants agreed to be interviewed at the end of the study. This could be an exception given our low participant numbers, but it could also be linked to the shame and fear of stigma reported by those who binge eat [5]. Exit interviews are an important part of the research process, therefore moving forwards we aim to explore the potential of questionnaire-based exit interviews that may be perceived by participants as more confidential.

Limitations

This was an exploratory study aiming to explore the type of mobile sensing data that might be relevant in detecting episodes of binge eating. We acknowledge the limitations of our small sample size (20 participants) which makes it difficult to conclude any definitive findings, particularly given the individual differences between participants. However, this preliminary study was conducted with a view towards informing the design of future larger scale trials in this space, and given what we now know about participants' engagement with the study, their acceptance of the methods and their willingness to provide self reported data, we are confident that a larger scale study would be feasible. In addition, our study monitored a UK-only cohort of people with binge eating behaviours, limiting our results' generalizability to other contexts. During the study, our participants' general behaviour might not be representative of their usual routines because of mobility limitations during COVID's lockdown. Our smartphone monitoring app was only compatible with Android and stopped collecting sensor data in four of our participants' phones, likely due to a software bug related to data synchronisation; that, as it is, could limit the deployment of future studies using the same app. The exploratory nature of our study called for the collection of data from multiple smartphone sensors. However, this might have influenced our low initial consent rate, compared to the number of people initially interested in participating. Clearer study information materials provided early in the recruitment process and a more constrained sensing approach might alleviate this limitation.

Conclusions

We shared a preliminary analysis of the differences in smartphone-based behavioural features between days with and without binge eating episodes to explore the feasibility of using mobile sensing to detect these events. We contextualised the experiences of people that binge eat and reflected on the challenges and opportunities of working with this population. Additionally, we discussed the need to understand participants' personal and social contexts preceding and accompanying their binge eating episodes, to be able to weigh the benefits, constraints and risks of monitoring them using smartphones, as well as the implications of leveraging the insights extracted from these data sources to plan for safer and more effective digital interventions.

Acknowledgements

All authors contributed to the preparation of the manuscript. Analysis of quantitative data was conducted by Julio Vega and Jue Xie; analysis of qualitative data was carried out by Heidi Ng and Roisin McNaney; the study was carried out by Beth Bell and Caitlin Taylor. The study was funded through a UK EPSRC-NIHR NewMinds Network plus grant. We thank Sarah Clinch (University of Manchester, UK), Aaron Sefi (Kooth, UK) and Leo Kroll (clinical psychologist) for their early input into the study design.

Conflicts of Interest

None declared

Abbreviations

App: Application
BMI: Body Mass Index
CBT: Cognitive Behavioural Therapy
COVID-19: coronavirus disease
DeMMI app: Detecting Mental health behaviours using Mobile Interactions application

Multimedia Appendices

1. What You Do/ What We See resource
2. Installation instructions provided to participants

References

1. Honary M, Bell BT, Clinch S, Wild SE, McNaney R. Understanding the Role of Healthy Eating and Fitness Mobile Apps in the Formation of Maladaptive Eating and Exercise Behaviors in Young People. JMIR mHealth and uHealth 2019;7(6):e14239. [doi: 10.2196/14239]
2. Wick MR, Fitzgerald EH, Keel PK. Epidemiology of Binge Eating. In: Frank GKW, Berner LA, editors. Binge Eating: A Transdiagnostic Psychopathology [Internet] Cham: Springer International Publishing; 2020 [cited 2020 Sep 15]. p. 3–12. [doi: 10.1007/978-3-030-43562-2_1]
3. Diagnostic and statistical manual of mental disorders. 3d ed. Washington, D.C: American Psychiatric Association; 1980.
4. Bulimia nervosa - Symptoms and causes [Internet]. Mayo Clinic. [cited 2020 Sep 15]. Available from: https://www.mayoclinic.org/diseases-conditions/bulimia/symptoms-causes/syc-20353615
5. Hepworth N, Paxton SJ. Pathways to help-seeking in bulimia nervosa and binge eating problems: A concept mapping approach. The International journal of eating disorders Hoboken: Wiley; 2007;40(6):493–504. [doi: 10.1002/eat.20402]
6. Swan S, Andrews B. The relationship between shame, eating disorders and disclosure in treatment. British Journal of Clinical Psychology 2003;42(4):367–378. [doi: 10.1348/014466503322528919]
7. Jansen A, Van den Hout M, Griez E. Clinical and non-clinical binges. Behaviour Research and Therapy Netherlands: Elsevier Science; 1990;28(5):439–444. [doi: 10.1016/0005-7967(90)90164-E]
8. Ivezaj V, Saules KK, Hoodin F, Alschuler K, Angelella NE, Collings AS, Saunders-Scott D, Wiedemann AA. The relationship between binge eating and weight status on depression, anxiety, and body image among a diverse college sample: A focus on Bi/Multiracial women. Eating Behaviors 2010 Jan 1;11(1):18–24. [doi: 10.1016/j.eatbeh.2009.08.003]
9. Abebe DS, Lien L, Torgersen L, von Soest T. Binge eating, purging and non-purging compensatory behaviours decrease from adolescence to adulthood: A population-based, longitudinal study. BMC Public Health 2012 Jan 13;12(1):32. [doi: 10.1186/1471-2458-12-32]
10. Schulz S, Laessle RG. Associations of negative affect and eating behaviour in obese women with and without binge eating disorder. Eat Weight Disord 2010 Dec 1;15(4):e287–e293. [doi: 10.1007/BF03325311]
11. Gross JJ, Uusberg H, Uusberg A. Mental illness and well-being: an affect regulation perspective. World Psychiatry 2019;18(2):130–139. [doi:

10.1002/wps.20618]
12. Dingemans AE, Furth EF van. Binge eating disorder psychopathology in normal weight and obese individuals. International Journal of Eating Disorders 2012;45(1):135–138. [doi: 10.1002/eat.20905]
13. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J-P. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology Nature Publishing Group; 2018 Jul;43(8):1660–1666. [doi: 10.1038/s41386-018-0030-z]
14. Matthews M, Abdullah S, Murnane E, Voida S, Choudhury T, Gay G, Frank E. Development and Evaluation of a Smartphone-Based Measure of Social Rhythms for Bipolar Disorder. Assessment SAGE Publications Inc; 2016 Aug 1;23(4):472–483. [doi: 10.1177/1073191116656794]
15. Trajectories of depression | Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing [Internet]. [cited 2020 Sep 16]. Available from: https://dl.acm.org/doi/abs/10.1145/2750858.2805845?casa_token=4I_Dm9JUdQcAAAAA:-QBbxolce6mlNFzZzmtG_6tw_juveGM4tMwzaQKCdMP5GVuSbbEd-GKu2kvBRtUOVLzS1fYZCgst
16. Juarascio AS, Crochiere RJ, Tapera TM, Palermo M, Zhang F. Momentary changes in heart rate variability can detect risk for emotional eating episodes. Appetite 2020 Sep 1;152:104698. [doi: 10.1016/j.appet.2020.104698]
17. Juarascio AS, Parker MN, Lagacey MA, Godfrey KM. Just-in-time adaptive interventions: A novel approach for enhancing skill utilization and acquisition in cognitive behavioral therapy for eating disorders. International Journal of Eating Disorders 2018;51(8):826–830. [doi: 10.1002/eat.22924]
18. Juarascio AS, Manasse SM, Goldstein SP, Forman EM, Butryn ML. Review of Smartphone Applications for the Treatment of Eating Disorders. European Eating Disorders Review 2015;23(1):1–11. [doi: 10.1002/erv.2327]
19. Burton AL, Abbott MJ. Conceptualising Binge Eating: A Review of the Theoretical and Empirical Literature. Behaviour Change Cambridge University Press; 2017 Sep;34(3):168–198. [doi: 10.1017/bec.2017.12]
20. Ambwani S, Roche MJ, Minnick AM, Pincus AL. Negative affect, interpersonal perception, and binge eating behavior: An experience sampling study. International Journal of Eating Disorders Wiley Online Library; 2015;48(6):715–726.
21. Goldschmidt AB, Crosby RD, Engel SG, Crow SJ, Cao L, Peterson CB, Durkin N. Affect and eating behavior in obese adults with and without elevated depression symptoms. International Journal of Eating Disorders Wiley Online Library; 2014;47(3):281–286.
22. Hilbert A, Tuschen-Caffier B. Maintenance of binge eating through negative mood: A naturalistic comparison of binge eating disorder and bulimia nervosa. International Journal of Eating Disorders Wiley Online Library; 2007;40(6):521–530.
23. Smyth JM, Wonderlich SA, Heron KE, Sliwinski MJ, Crosby RD, Mitchell JE, Engel SG. Daily and momentary mood and stress are associated with binge eating and vomiting in bulimia nervosa patients in the natural environment. Journal of

consulting and clinical psychology American Psychological Association; 2007;75(4):629.
24. Andrewes HE, Hulbert C, Cotton SM, Betts J, Chanen AM. An ecological momentary assessment investigation of complex and conflicting emotions in youth with borderline personality disorder. Psychiatry Research Elsevier; 2017;252:102–110.
25. Houben M, Claes L, Vansteelandt K, Berens A, Sleuwaegen E, Kuppens P. The emotion regulation function of nonsuicidal self-injury: A momentary assessment study in inpatients with borderline personality disorder features. Journal of Abnormal Psychology American Psychological Association; 2017;126(1):89.
26. Khue LM, Ouh EL, Jarzabek S. Mood self-assessment on smartphones. Proceedings of the conference on Wireless Health 2015. p. 1–8.
27. Lathia N, Sandstrom GM, Mascolo C, Rentfrow PJ. Happier people live more active lives: Using smartphones to link happiness and physical activity. PloS one Public Library of Science San Francisco, CA USA; 2017;12(1):e0160589.
28. Panza E. Minority stress and eating behavior among overweight and obese sexual minority women: an ecological momentary assessment study [PhD Thesis]. Rutgers University-School of Graduate Studies; 2018.
29. Santangelo PS, Koenig J, Funke V, Parzer P, Resch F, Ebner-Priemer UW, Kaess M. Ecological momentary assessment of affective and interpersonal instability in adolescent non-suicidal self-injury. Journal of abnormal child psychology Springer; 2017;45(7):1429–1438.
30. Schwartz S, Schultz S, Reider A, Saunders EF. Daily mood monitoring of symptoms using smartphones in bipolar disorder: a pilot study assessing the feasibility of ecological momentary assessment. Journal of Affective Disorders Elsevier; 2016;191:88–93.
31. Pejovic V, Lathia N, Mascolo C, Musolesi M. Mobile-based experience sampling for behaviour research. Emotions and personality in personalized services Springer; 2016. p. 141–161.
32. Dubad M, Winsper C, Meyer C, Livanou M, Marwaha S. A systematic review of the psychometric properties, usability and clinical impacts of mobile mood-monitoring applications in young people. Psychological medicine Cambridge University Press; 2018;48(2):208–228.
33. Kanjo E, Kuss DJ, Ang CS. NotiMind: utilizing responses to smart phone notifications as affective sensors. IEEE Access IEEE; 2017;5:22023–22035.
34. Stieger S, Lewetz D. A week without using social media: Results from an ecological momentary intervention study using smartphones. Cyberpsychology, Behavior, and Social Networking Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New …; 2018;21(10):618–624.
35. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of personality and social psychology American Psychological Association; 1988;54(6):1063.
36. Pollak JP, Adams P, Gay G. PAM: a photographic affect meter for frequent, in situ measurement of affect. Proceedings of the SIGCHI conference on Human factors in computing systems 2011. p. 725–734.

37. Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry Pergamon; 1994;25(1):49–59.
38. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, Zhou X, Ben-Zeev D, Campbell AT. StudentLife: Using smartphones to assess mental health and academic performance of college students. Mobile health Springer; 2017. p. 7–33.
39. Griffin B, Saunders KEA. Smartphones and Wearables as a Method for Understanding Symptom Mechanisms. Front Psychiatry [Internet] 2020 Jan 17 [cited 2020 Sep 15];10. PMID:32009990
40. Onnela J-P, Rauch SL. Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. Neuropsychopharmacology Nature Publishing Group; 2016 Jun;41(7):1691–1696. [doi: 10.1038/npp.2016.7]
41. Aledavood T, Torous J, Triana Hoyos AM, Naslund JA, Onnela J-P, Keshavan M. Smartphone-Based Tracking of Sleep in Depression, Anxiety, and Psychotic Disorders. Curr Psychiatry Rep 2019 Jun 4;21(7):49. [doi: 10.1007/s11920-019-1043-y]
42. Masud MT, Mamun MA, Thapa K, Lee DH, Griffiths MD, Yang S-H. Unobtrusive monitoring of behavior and movement patterns to detect clinical depression severity level via smartphone. Journal of Biomedical Informatics 2020 Mar 1;103:103371. [doi: 10.1016/j.jbi.2019.103371]
43. Mehrotra A, Hendley R, Musolesi M. Towards multi-modal anticipatory monitoring of depressive states through the analysis of human-smartphone interaction. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct [Internet] New York, NY, USA: Association for Computing Machinery; 2016 [cited 2020 Sep 16]. p. 1132–1138. [doi: 10.1145/2968219.2968299]
44. Asare KO, Terhorst Y, Vega J, Peltonen E, Lagerspetz E, Ferreira D. Predicting Depression From Smartphone Behavioral Markers Using Machine Learning Methods, Hyperparameter Optimization, and Feature Importance Analysis: Exploratory Study. JMIR mHealth and uHealth 2021 Jul 12;9(7):e26540. [doi: 10.2196/26540]
45. Ben-Zeev D, Wang R, Abdullah S, Brian R, Scherer EA, Mistler LA, Hauser M, Kane JM, Campbell A, Choudhury T. Mobile Behavioral Sensing for Outpatients and Inpatients With Schizophrenia. PS American Psychiatric Publishing; 2015 Dec 15;67(5):558–561. [doi: 10.1176/appi.ps.201500130]
46. Buck B, Scherer E, Brian R, Wang R, Wang W, Campbell A, Choudhury T, Hauser M, Kane JM, Ben-Zeev D. Relationships between smartphone social behavior and relapse in schizophrenia: A preliminary report. Schizophrenia Research 2019 Jun 1;208:167–172. [doi: 10.1016/j.schres.2019.03.014]
47. Staples P, Torous J, Barnett I, Carlson K, Sandoval L, Keshavan M, Onnela J-P. A comparison of passive and active estimates of sleep in a cohort with schizophrenia. npj Schizophrenia Nature Publishing Group; 2017 Oct 16;3(1):1–6. [doi: 10.1038/s41537-017-0038-0]
48. Wang R, Wang W, Aung MSH, Ben-Zeev D, Brian R, Campbell AT, Choudhury T, Hauser M, Kane J, Scherer EA, Walsh M. Predicting Symptom Trajectories of Schizophrenia using Mobile Sensing. Proc ACM Interact Mob Wearable

Ubiquitous Technol 2017 Sep 11;1(3):110:1-110:24. [doi: 10.1145/3130976]
49. Alvarez-Lozano J, Osmani V, Mayora O, Frost M, Bardram J, Faurholt-Jepsen M, Kessing LV. Tell me your apps and I will tell you your mood: correlation of apps usage with bipolar disorder state. Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments [Internet] New York, NY, USA: Association for Computing Machinery; 2014 [cited 2020 Sep 16]. p. 1–7. [doi: 10.1145/2674396.2674408]
50. Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G, Rössler W. Using Smartphones to Monitor Bipolar Disorder Symptoms: A Pilot Study. JMIR Mental Health 2016;3(1):e2. [doi: 10.2196/mental.4560]
51. Carr O, Andreotti F, Saunders KEA, Palmius N, Goodwin GM, De Vos M. Monitoring Depression in Bipolar Disorder using Circadian Measures from Smartphone Accelerometers. arXiv:200702064 [cs, eess] [Internet] 2020 Jul 4 [cited 2020 Sep 17]; Available from: http://arxiv.org/abs/2007.02064
52. Faurholt-Jepsen M, Bauer M, Kessing LV. Smartphone-based objective monitoring in bipolar disorder: status and considerations. International Journal of Bipolar Disorders 2018 Jan 23;6(1):6. [doi: 10.1186/s40345-017-0110-8]
53. Puiatti A, Mudda S, Giordano S, Mayora O. Smartphone-centred wearable sensors network for monitoring patients with bipolar disorder. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2011. p. 3644–3647. [doi: 10.1109/IEMBS.2011.6090613]
54. Bauer G, Lukowicz P. Can smartphones detect stress-related changes in the behaviour of individuals? 2012 IEEE International Conference on Pervasive Computing and Communications Workshops 2012. p. 423–426. [doi: 10.1109/PerComW.2012.6197525]
55. Ciman M, Wac K, Gaggi O. iSensestress: Assessing stress through human-smartphone interaction analysis. 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) 2015. p. 84–91. [doi: 10.4108/icst.pervasivehealth.2015.259280]
56. Ferdous R, Osmani V, Mayora O. Smartphone app usage as a predictor of perceived stress levels at workplace. 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) 2015. p. 225–228. [doi: 10.4108/icst.pervasivehealth.2015.260192]
57. Garcia-Ceja E, Osmani V, Mayora O. Automatic Stress Detection in Working Environments From Smartphones' Accelerometer Data: A First Step. IEEE Journal of Biomedical and Health Informatics 2016 Jul;20(4):1053–1060. [doi: 10.1109/JBHI.2015.2446195]
58. Mayya S, Jilla V, Tiwari VN, Nayak MM, Narayanan R. Continuous monitoring of stress on smartphone using heart rate variability. 2015 IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE) 2015. p. 1–5. [doi: 10.1109/BIBE.2015.7367627]
59. Boukhechba M, Huang Y, Chow P, Fua K, Teachman BA, Barnes LE. Monitoring social anxiety from mobility and communication patterns. Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers [Internet] New York, NY, USA: Association for Computing

Machinery; 2017 [cited 2020 Sep 16]. p. 749–753. [doi: 10.1145/3123024.3125607]
60. Fukazawa Y, Ito T, Okimura T, Yamashita Y, Maeda T, Ota J. Predicting anxiety state using smartphone-based passive sensing. Journal of Biomedical Informatics 2019 May 1;93:103151. [doi: 10.1016/j.jbi.2019.103151]
61. Gong J, Huang Y, Chow PI, Fua K, Gerber MS, Teachman BA, Barnes LE. Understanding behavioral dynamics of social anxiety among college students through smartphone sensors. Information Fusion 2019 Sep 1;49:57–68. [doi: 10.1016/j.inffus.2018.09.002]
62. Rashid H, Mendu S, Daniel KE, Beltzer ML, Teachman BA, Boukhechba M, Barnes LE. Predicting Subjective Measures of Social Anxiety from Sparsely Collected Mobile Sensor Data. Proc ACM Interact Mob Wearable Ubiquitous Technol 2020 Sep 4;4(3):109:1-109:24. [doi: 10.1145/3411823]
63. Berry N, Bucci S, Lobban F. Use of the Internet and Mobile Phones for Self-Management of Severe Mental Health Problems: Qualitative Study of Staff Views. JMIR Mental Health 2017;4(4):e52. [doi: 10.2196/mental.8311]
64. Fairburn CG, Rothwell ER. Apps and eating disorders: A systematic clinical appraisal. International Journal of Eating Disorders 2015;48(7):1038–1046. [doi: 10.1002/eat.22398]
65. Lindgreen P, Lomborg K, Clausen L. Patient Experiences Using a Self-Monitoring App in Eating Disorder Treatment: Qualitative Study. JMIR mHealth and uHealth 2018 Jun 22;6(6):e10253. [doi: 10.2196/10253]
66. Recovery Record - Best Practice for Eating Disorders Treatment [Internet]. [cited 2020 Sep 17]. Available from: https://www.recoveryrecord.com/
67. Tulu B, Ruiz C, Allard J, Acheson J, Busch A, Roskuski A, Heeringa G, Jaskula V, Oleski J, Pagoto S. SlipBuddy: A mobile health intervention to prevent overeating. Proceedings of the 50th Hawaii international conference on system sciences 2017.
68. Ferreira D, Kostakos V, Dey AK. AWARE: Mobile Context Instrumentation Framework. Front ICT [Internet] Frontiers; 2015 [cited 2020 Jul 24];2. [doi: 10.3389/fict.2015.00006]
69. Branley-Bell D, Talbot CV. Exploring the impact of the COVID-19 pandemic and UK lockdown on individuals with experience of eating disorders. PsyArXiv; 2020;
70. Schlegl S, Maier J, Meule A, Voderholzer U. Eating disorders in times of the COVID-19 pandemic—Results from an online survey of patients with anorexia nervosa. International Journal of Eating Disorders Wiley Online Library; 2020;
71. Vega J. RAPIDS [Internet]. Reproducible Analysis Pipeline for Data Streams. 2021 [cited 2020 Nov 18]. Available from: https://www.rapids.science/latest/
72. Vega J, Li M, Aguillera K, Goel N, Joshi E, Khandekar K, Durica KC, Kunta AR, Low CA. Reproducible Analysis Pipeline for Data Streams: Open-Source Software to Process Data Collected With Mobile Devices. Frontiers in Digital Health [Internet] 2021 [cited 2022 Jan 16];3. Available from: https://www.frontiersin.org/article/10.3389/fdgth.2021.769823
73. Parker RI, Vannest K. An improved effect size for single-case research: nonoverlap of all pairs. Behav Ther 2009 Dec;40(4):357–367. PMID:19892081

74. Elo S, Kyngäs H. The qualitative content analysis process. Journal of Advanced Nursing 2008;62(1):107–115. [doi: 10.1111/j.1365-2648.2007.04569.x]
75. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology Routledge; 2006 Jan 1;3(2):77–101. [doi: 10.1191/1478088706qp063oa]
76. JulioV. JulioV/demmi_aware [Internet]. 2021 [cited 2021 Jul 7]. Available from: https://github.com/JulioV/demmi_aware
77. JulioV. JulioV/demmi-analysis [Internet]. 2021 [cited 2021 Jul 7]. Available from: https://github.com/JulioV/demmi-analysis
78. Pustejovsky JE, Swan DM. SingleCaseES: A Calculator for Single-Case Effect Sizes [Internet]. 2019. Available from: https://CRAN.R-project.org/package=SingleCaseES
79. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) [Royal Statistical Society, Wiley]; 1995;57(1):289–300.
80. Haedt-Matt AA, Keel PK. Revisiting the affect regulation model of binge eating: A meta-analysis of studies using ecological momentary assessment. Psychological Bulletin US: American Psychological Association; 2011;137(4):660–681. [doi: 10.1037/a0023660]
81. Jacobson NC, Chung YJ. Passive Sensing of Prediction of Moment-To-Moment Depressed Mood among Undergraduates with Clinical Levels of Depression Sample Using Smartphones. Sensors Multidisciplinary Digital Publishing Institute; 2020 Jan;20(12):3572. [doi: 10.3390/s20123572]
82. Dingemans A, Danner U, Parks M. Emotion Regulation in Binge Eating Disorder: A Review. Nutrients Multidisciplinary Digital Publishing Institute; 2017 Nov;9(11):1274. [doi: 10.3390/nu9111274]
83. Wilson GT, Sysko R. Frequency of Binge Eating Episodes in Bulimia Nervosa and Binge Eating Disorder: Diagnostic Considerations. Int J Eat Disord [Internet] 2009 Nov [cited 2020 Sep 16];42(7). PMID:19610014
84. Rossiter EM, Agras WS, Telch CF, Bruce B. The eating patterns of non-purging bulimic subjects. International Journal of Eating Disorders 1992;11(2):111–120. [doi: 10.1002/1098-108X(199203)11:2<111::AID-EAT2260110203>3.0.CO;2-J]
85. Fellger A, Sprint G, Weeks D, Crooks E, Cook DJ. Wearable Device-Independent Next Day Activity and Next Night Sleep Prediction for Rehabilitation Populations. IEEE Journal of Translational Engineering in Health and Medicine 2020;8:1–9. [doi: 10.1109/JTEHM.2020.3014564]
86. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of Smartphone Applications and Wearable Devices for Tracking Physical Activity Data. JAMA 2015 Feb 10;313(6):625–626. [doi: 10.1001/jama.2014.17841]
87. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The Effect of Physical Activity Interventions Comprising Wearables and Smartphone Applications on Physical Activity: a Systematic Review and Meta-analysis. Sports Med - Open 2018 Sep 3;4(1):42. [doi: 10.1186/s40798-018-0157-9]
88. Hilbert A, Petroff D, Herpertz S, Pietrowsky R, Tuschen-Caffier B, Vocks S,

Schmidt R. Meta-analysis of the efficacy of psychological and medical treatments for binge-eating disorder. Journal of Consulting and Clinical Psychology US: American Psychological Association; 2019;87(1):91–105. [doi: 10.1037/ccp0000358]
89. Linardon J, Wade TD. How many individuals achieve symptom abstinence following psychological treatments for bulimia nervosa? A meta-analytic review. International Journal of Eating Disorders 2018;51(4):287–294. [doi: 10.1002/eat.22838]
90. Linardon J, Wade TD, de la Piedad Garcia X, Brennan L. The efficacy of cognitive-behavioral therapy for eating disorders: A systematic review and meta-analysis. Journal of Consulting and Clinical Psychology US: American Psychological Association; 2017;85(11):1080–1094. [doi: 10.1037/ccp0000245]
91. Honary M, Bell B, Clinch S, Vega J, Kroll L, Sefi A, McNaney R. Shaping the Design of Smartphone-Based Interventions for Self-Harm. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems [Internet] New York, NY, USA: Association for Computing Machinery; 2020 [cited 2020 Sep 16]. p. 1–14. [doi: 10.1145/3313831.3376370]
92. Cohen DA, Wu S-Y, Farley TA. Comparing the Cost-Effectiveness of HIV Prevention Interventions. JAIDS Journal of Acquired Immune Deficiency Syndromes 2004 Nov 1;37(3):1404–1414. [doi: 10.1097/01.qai.0000123271.76723.96]
93. Cobiac LJ, Vos T, Barendregt JJ. Cost-Effectiveness of Interventions to Promote Physical Activity: A Modelling Study. PLOS Medicine Public Library of Science; 2009 Jul 14;6(7):e1000110. [doi: 10.1371/journal.pmed.1000110]
94. Tsevat J. Impact and cost-effectiveness of smoking interventions. The American Journal of Medicine 1992 Jul 15;93(1, Supplement 1):S43–S47. [doi: 10.1016/0002-9343(92)90627-N]
95. Wutzke SE, Shiell A, Gomel MK, Conigrave KM. Cost effectiveness of brief interventions for reducing alcohol consumption. Social Science & Medicine 2001 Mar 1;52(6):863–870. [doi: 10.1016/S0277-9536(00)00189-1]
96. Lynch FL, Striegel-Moore RH, Dickerson JF, Perrin N, DeBar L, Wilson GT, Kraemer HC. Cost-effectiveness of guided self-help treatment for recurrent binge eating. Journal of Consulting and Clinical Psychology US: American Psychological Association; 2010;78(3):322–333. [doi: 10.1037/a0018982]
97. Reychav I, Beeri R, Balapour A, Raban DR, Sabherwal R, Azuri J. How reliable are self-assessments using mobile technology in healthcare? The effects of technology identity and self-efficacy. Computers in Human Behavior 2019 Feb 1;91:52–61. [doi: 10.1016/j.chb.2018.09.024]
98. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. Controlled Clinical Trials 2003 Apr 1;24(2):182–199. [doi: 10.1016/S0197-2456(02)00320-3]
99. Swan M. The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. Big Data Mary Ann Liebert, Inc., publishers; 2013 Jun 1;1(2):85–99. [doi: 10.1089/big.2012.0002]
100. Or CKL, Karsh B-T. A Systematic Review of Patient Acceptance of Consumer

Health Information Technology. J Am Med Inform Assoc 2009;16(4):550–560. PMID:19390112
101. Or C, Tao D. Usability Study of a Computer-Based Self-Management System for Older Adults with Chronic Diseases. JMIR Research Protocols 2012 Nov 8;1(2):e2184. [doi: 10.2196/resprot.2184]
102. Nielsen J. Engineering, Usability. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA; 1993.
103. Honary M, Fisher NR, McNaney R, Lobban F. A Web-Based Intervention for Relatives of People Experiencing Psychosis or Bipolar Disorder: Design Study Using a User-Centered Approach. JMIR Mental Health 2018 Dec 7;5(4):e11473. [doi: 10.2196/11473]