

Student Perspectives on Digital Phenotyping

The Acceptability of Using Smartphone Data to Assess Mental Health

John Rooksby

Northumbria University
Newcastle-upon-Tyne, UK
john.rooksby@northumbria.ac.uk

Alistair Morrison

University of Glasgow
Glasgow, UK
alistair.morrison@glasgow.ac.uk

Dave Murray-Rust

University of Edinburgh
Edinburgh, UK
d.murray-rust@ed.ac.uk

ABSTRACT

There is a mental health crisis facing universities internationally. A growing body of interdisciplinary research has successfully demonstrated that using sensor and interaction data from students' smartphones can give insight into stress, depression, mood, suicide risk and more. The approach, which is sometimes termed Digital Phenotyping, has potential to transform how mental health and wellbeing can be monitored and understood. The approach could also transform how interventions are designed, delivered and evaluated. To date, little work has addressed the human and ethical side of digital phenotyping, including how students feel about being monitored. In this paper we report findings from in-depth focus groups, prototyping and interviews with students. We find they are positive about mental health technology, but also that there are multi-layered issues to address if digital phenotyping is to become acceptable. Using an acceptability framework, we set out the key design challenges that need to be addressed.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**;

KEYWORDS

Qualitative Research; Acceptability; Mobile Health; Mental Health; Mental Wellbeing; Lived Informatics; Sensors.

ACM Reference Format:

John Rooksby, Alistair Morrison, and Dave Murray-Rust. 2019. Student Perspectives on Digital Phenotyping: The Acceptability of

Using Smartphone Data to Assess Mental Health. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, May 4–9, 2019, Glasgow, Scotland UK. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3290605.3300655>

1 INTRODUCTION

Student mental health and wellbeing is of concern internationally. Mental health and emotional wellbeing among college students in the USA is in "*continued decline*" [25, 82], an increasing suicide rate in South Korea is the "*hidden price of education*" [38], and in Australia there is an "*urgent need to better understand the prevalence of mental health problems*" among tertiary students [66].

From a UK perspective, mental health problems are increasing among students and young people, and the large majority of students will experience some form of emotional distress [49, 80, 83]. One prominent cause for concern has been an increasing suicide rate [32]. However, there is a much wider problem—students that experience mental illness and mental distress have a lower quality of life, achieve less, and are more likely to drop out from education [16, 58, 80].

Universities have a duty of care to students, and many offer counselling and support services. However, these services are increasingly stretched and many students feel under-supported [11]. Services themselves are turning to new forms of limited counselling and to online services in order to meet demand [15]. A report by the UK Institute for Public Policy Research has recommended: "*Universities should not just be helping people in crisis but also concentrating on prevention, early intervention, management of risk and giving low level support*" [80].

In this paper we are interested in an emerging form of health surveillance technology, sometimes referred to as *Digital Phenotyping*, that uses passive sensing to identify and monitor problems. Data is collected via smartphones [60, 84], social media [5, 70], wearables [52], eLearning platforms [67] and more, which may then be useful for: (i) Monitoring students known to be at risk or with pre-diagnosed disorders and self reported problems; (ii) Monitoring all students to identify individuals who may be at risk and requiring help; (iii) Monitoring the student body as a whole in order to measure mental health and wellbeing and inform policy and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-5970-2/19/05...\$15.00

<https://doi.org/10.1145/3290605.3300655>

Table 1: Smartphone sensors previously used in digital phenotyping research

| Sensor | Description |
|-----------------|--------------------------------------|
| Accel./Activity | Movement of device and person |
| App Use | App launches, installs, etc. |
| Web History | Websites visited |
| Battery | Battery level and charging |
| Bluetooth | Devices seen, plus status |
| Call Logs | Calls made and received |
| Camera | Raw images, num photos taken, etc. |
| Screen | Screen status (on/off) |
| Keyboard/UI | Event counts, potentially keylogging |
| Location | Geographical coordinates of device |
| Light | Light levels detected |
| Microphone | Sound recordings, decibels, etc. |
| SMS/Email | Messages sent and received |

Table 2: Self reports previously used in digital phenotyping research

| Survey | Description |
|--------|---|
| PHQ9 | Depression and low mood scale (see [43]) |
| GAD7 | Generalised Anxiety Disorder Scale (see [77]) |
| WEMWBS | Mental Wellbeing Scale (see [79]) |

service provision; (iv) Monitoring students solely for research purposes.

As this area matures and expands, it is timely to explore students' perspectives on the acceptability of digital phenotyping systems for monitoring, assessing and researching mental health and wellbeing. Acceptability is clearly crucial if the technology is to be adopted beyond limited studies. In this paper we will specifically explore the acceptability of *smartphone-based* digital phenotyping. Smartphones are the source for a large part of research in this area, including technologies for students.

2 BACKGROUND

Dawkins [24] argued that the Life Sciences should examine the behavioural expressions of organisms, or what he called "*the extended phenotype*". This idea has found traction in areas including Psychiatry, where the work of Jain et al. [40] sparked interest in behavioural expressions of mental health and wellbeing. Jain et al. explain that collecting behavioural data in mental health allows for "*a more comprehensive and nuanced view of the experience of illness*", providing "*substantial value above and beyond the physical exam, laboratory values and clinical imaging data—our traditional approaches to characterizing a disease phenotype*" [40].

Digital phenotyping and smartphone data

In Psychiatry, the term *digital phenotyping* refers to the use of digital technology to measure the extended phenotype [39]. Jain et al. describe the use of interaction data from smartphones and computers, data from wearable sensors, web browsing and search data, and social media data [40]. Many place particular value on smartphone data; for example Torous et al. state "*The data generated by increasingly sophisticated smartphone sensors and phone use patterns appear ideal for capturing various social and behavioral dimensions of psychiatric and neurological diseases*" [81]. Onnela et al. discuss digital phenotyping with special reference to smartphone data, defining it as: "*the moment-by-moment quantification of the individual-level human phenotype in situ using data from personal digital devices, in particular smartphones*" [62]. One study has even coined the term "*phone-o-typing*" [73].

Work to date in this area has focused on several issues including monitoring people already diagnosed with mental health issues such as schizophrenia [7, 10, 86, 87] and depression [17, 18], and monitoring general populations for signs of depression [69] or examining mood [55]. Other work has looked at specific populations and contexts, for example looking at stress in the workplace [29]. Much of the current work has been of limited scale and primarily for research, but wide-scale monitoring is envisioned in this area, such as the systematic, population scale data collections proposed by [2, 50, 75, 81] and others.

Data collected. Typical forms of data collected in digital phenotyping are reported in Tables 1 and 2. The key form in this area is sensor and interaction data (Table 1). These data are generally 'passive' in that they are not actively input by the user, but are generated incidentally during day-to-day life. Unlike much of the personal tracking data discussed in HCI and contexts of the quantified self [59, 68], this data can be more 'raw' in nature rather than something intended to hold meaning to end users, with, for example, accelerometer logs often being collected rather than step counts.

Another important form of data are self-reports. In Table 2, we summarise several standard questionnaires that are often used. These are medical questionnaires and not designed specifically for self-tracking. The point of such questionnaires is often for providing 'ground truth' about the person, with which to then compare with the sensor data. These questionnaires are important for research, but the vision for digital phenotyping is primarily that passive data can be enough to inform assessments and meaningful monitoring. Other forms of self-report data can include demographics, medication logs and schedules or timetables. Several systems have also used Experience Sampling Methods (ESM) [19, 35, 85] in which self-report questions are triggered at

various points in time or associated with contextual factors such as location.

Digital phenotyping and students

A population of particular interest in digital phenotyping research has been students. The work with students has often been exploratory, collecting data for many purposes, as summarised in Table 3. The table includes the focus (i.e. what the research was monitoring or attempting to infer), and descriptions of what ‘raw’ data was collected. The table does not show sampling rates, consider how models were constructed, or consider findings— but is intended to characterise the scope and direction of the area.

A widely discussed study of students is the StudentLife project by Wang et al. [84]. The authors developed a “*continuous sensing app*” that collected a variety of data, as detailed in Table 3. They used this data in order to assess the day-by-day and week-by-week impact of student life on stress, sleep, activity, mood, sociability, mental wellbeing and academic performance. The results from the study are insightful, showing correlations between the data with mental health and educational outcomes.

Data from the StudentLife study was collected from students participating in a computer science class. They were given mobile devices to use, preferably as their primary device. The authors gained ethical approval for the study and gained consent from the students, but it is not clear how the students felt about being tracked by their university, and it is not clear what opportunities they were given to talk through their thoughts. The authors mention gaining consent from participants but in the context of students being “*trained to use the app*” and shown “*how to respond to the MobileEMA system*”. Similar accounts are given in many other papers from Table 3, where contact with participants is about ensuring they comply with research procedures in order that data of sufficient quality may be gained in order to perform valid analyses. Ensuring compliance is important and reasonable in early stage research, but does not wholly enable exploration of autonomy and acceptability at the same time.

Perspectives in Human Computer Interaction (HCI). The StudentLife study has been influential in HCI research, and has been prominently discussed in papers by Kelley et al. [41] and Lee and Hong [44]. Kelley et al. consider the ways in which tracked data might be put to use by counsellors, and Lee and Hong consider the development of personalised mental health interventions by students. This perspective is characteristic of the wider body of work on activity tracking in HCI which has been human centered but predominantly concerned with facilitating uptake of tracking technology. An exception is Mathews et al. [53], who have called for caution and critical perspectives in mental health tracking.

The literature on workplace tracking [6, 22, 65] is also more cautious. Our own work is highly influenced by Kelley et al. [41] and others, but with a more cautious tone.

Beyond HCI, Lovatt & Holmes [47] have critiqued digital phenotyping from a sociological perspective, praising the creation of new forms of measurement, but worrying it takes a reductive, individualistic stance on social behaviours.

Acceptability

Acceptability is an important consideration for health technologies and interventions [56, 71, 90]. If an intervention is acceptable then people are more likely to engage and adhere to it. Acceptability is not always the first consideration in intervention development (as has been the case with digital phenotyping), with efficacy often given greater initial priority. Acceptability, however, is an important dimension of effectiveness and one that ought to be addressed early. It also has an interrelationship with ethics, particularly concepts of autonomy and informed consent.

There are several ways in which acceptability can be defined. “Social acceptability” has been one consideration in HCI [61, 89] and health [72]. Another consideration has been the perspective of experts and those delivering technologies and interventions on acceptability; this perspective is evident in ethics research in HCI (e.g. [48, 54]). From the perspective of this paper, acceptability is what the user or beneficiary of the technology thinks and feels.

Sekhon et al. [71] have developed a Theoretical Framework for Acceptability (TFA) for health interventions. The framework centres on the user’s point of view, and is intended to be applied throughout the lifecycle of intervention development (prospectively, concurrently and retrospectively). We will draw from the TFA later in this paper for a prospective analysis of digital phenotyping.

Scope of this work

As an important note, the term *digital phenotyping* is not always used in the papers we describe here. For us, the term is primarily a way of referring to a growing body of work. However, use of the term also signifies that there are theoretical perspectives at play, ones that data science work does not always acknowledge. Similar perspectives include “*reality mining*” [26], “*social physics*” [64] or broader medical and sociological concepts of “*health surveillance*” [3, 34].

The work we are describing does have strong parallels with other forms of smartphone sensor-based study, particularly smartphone usage analytics [12, 23, 30, 57, 88], and a growing body of work on occupational stress (e.g. [29, 51]).

3 THE STUDY

We report on an in-depth study of the acceptability to students of digital phenotyping of mental health by universities.

Table 3: Papers using sensor data for monitoring student mental health

| Authors | Focus | Accel. / Activity | App usage | Battery / charge | Browser History | Bluetooth | Call logs | Camera events | Screen | Keyboard / UI | Location | Light Sensor | Microphone | SMS / Email | Duration | N |
|------------------------|----------------|-------------------|-----------|------------------|-----------------|-----------|-----------|---------------|--------|---------------|----------|--------------|------------|-------------|---------------|----|
| Abdullah et al. [1] | Sleep | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | | ✓ | 97 days | 9 |
| Asselbergs et al. [4] | Mood | ✓ | ✓ | | | | ✓ | ✓ | ✓ | | | | | ✓ | 35 days (avg) | 27 |
| Becker et al. [8] | Mood | ✓ | ✓ | | | | ✓ | | ✓ | | | | | ✓ | 6 weeks | 27 |
| Ben-Zeev et al. [9] | Multiple | ✓ | | | | | | | ✓ | | ✓ | ✓ | ✓ | | 10 weeks | 47 |
| Boukhechba et al. [13] | Social anxiety | | | | | | ✓ | | | | ✓ | | | ✓ | 2 weeks | 54 |
| Chan et al. [20] | Method / UI | ✓ | ✓ | | | | ✓ | | ✓ | | ✓ | | ✓ | ✓ | 12 days (avg) | 32 |
| Chen et al. [21] | Sleep | ✓ | | ✓ | | | | | ✓ | | | ✓ | ✓ | | 1 week | 8 |
| Eskes et al. [27] | Sociability | | ✓ | | | ✓ | ✓ | | | | ✓ | | | | 11 days (avg) | 10 |
| Farhan et al. [28] | Depression | ✓ | | | | | | | | | ✓ | | | | 14 day blocks | 79 |
| Huang et al. [36] | Social anxiety | | | | | | | | | | ✓ | | | | 10 days | 18 |
| Hung et al. [37] | Depression | | ✓ | | | | ✓ | | ✓ | | | | | | 14+10 days | 18 |
| Lee et al. [45] | Phone overuse | | ✓ | | ✓ | | ✓ | | ✓ | ✓ | | | | ✓ | 27 days (avg) | 95 |
| LiKamWa et al. [46] | Mood | | ✓ | | ✓ | | ✓ | | | | ✓ | | | ✓ | 2 months | 32 |
| Madan et al. [50] | General | | | | | ✓ | ✓ | | | | ✓ | | | ✓ | 2 months | 70 |
| Nobles et al. [60] | Suicide risk | | | | ✓ | | ✓ | | | | | | | ✓ | Historical | 26 |
| Singh et al. [73] | Cooperation | | | | | | ✓ | | | | ✓ | | | ✓ | 10 weeks | 54 |
| Singh et al. [74] | Social capital | | | | | | ✓ | | | | | | | ✓ | 10 weeks | 55 |
| Stütz et al. [78] | Stress | ✓ | ✓ | | | | ✓ | | ✓ | | | ✓ | ✓ | ✓ | 2 weeks | 15 |
| Wang et al. [85] | Multiple | ✓ | ✓ | | | ✓ | ✓ | | | | ✓ | ✓ | ✓ | ✓ | 10 weeks | 48 |

Our work has combined focus groups, interviews and deployment of a tracker application. We used focus groups to encourage groups of peers to engage in extended and open discussions, followed up with individual interviews to elicit personal opinions and reflection. Depth of discussion has been important in this work– we spent on average 4.5 hours (268.5 audio recorded minutes, min=236, max=316, sd=25) face-to-face with each student in group and individual sessions, where we discussed issues that were conceptually and sometimes emotionally difficult.

The study gained ethical approval from an IRB at the study site. We have used an "ongoing consent" approach [33], in which we gained informed consent at the outset, and returned to the study information throughout. At the end we showed participants their transcripts and log data for comment and discussion. Participants have also had the opportunity to comment on this paper. This was important because the students' views and understandings were formative during the study. Anonymisation in this paper is a little more stringent than usual because of participant's concerns about what they revealed.

Participants

15 students participated in our study (see Table 4). 11 participants were female and 4 male, 8 were undergraduate and 7 postgraduate. The average age was 23.5 (min=18, max=30, sd=3.6).

We did not recruit people based on whether they had experienced mental distress or mental health problems, our interest being in the general student population. Given that many students in the UK have experienced mental distress (80% of students experience stress and around 50% experience anxiety, problems sleeping and/or feelings of depression [42], 43% experience feelings of isolation and loneliness [58]) it was likely that people with these experiences would attend. Given that 28% of women in the UK have experienced mental health problems by the time they are 25 [80], it was also likely that we would have people with these experiences.

To preserve anonymity, specific ages, course information and device hardware are not given in Table 4. Participants were studying topics including history, architecture, design and art; none were studying computing or medical subjects. Participants were diverse, and included nationals of and/or

Table 4: Participant information for the three cohorts. Notes: (level: UG=undergraduate, M=masters), (times shown are total for focus groups + interview), (sensor data types: ✓ = participant’s personalised data recorded and visualised, ▲ =preprepared non-personal data visualisations shown, △=discussed only)

| | Age | Gender id. | Level | Device | Focus Groups | Interview | Time (FG+Int) | Screen | App Use | Battery | Location | Microphone | Activity | Bluetooth | Call logs | Camera | Keyboard | Light | SMS/Email | PHQ9 | GAD7 | WEMWBS |
|-----|-----|------------|-------|---------|--------------|-----------|---------------|--------|---------|---------|----------|------------|----------|-----------|-----------|--------|----------|-------|-----------|------|------|--------|
| P1 | ≥24 | f | M | Android | ✓✓ | ✓ | 5h16 | ✓ | ✓ | ✓ | ▲ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P2 | <24 | f | UG | Android | ✓✓ | ✓ | 5h03 | ✓ | ✓ | ✓ | ✓ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P3 | <24 | f | UG | iOS | ✓✓ | | 4h20 | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P4 | ≥24 | m | PhD | iOS | ✓✓ | ✓ | 5h06 | ✓ | ▲ | ✓ | ✓ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P5 | <24 | f | UG | iOS | ✓✓ | | 4h20 | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P6 | <24 | f | UG | Android | ✓✓ | ✓ | 4h29 | ✓ | ✓ | ✓ | ▲ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P7 | <24 | f | UG | iOS | ✓✓ | ✓ | 5h01 | ✓ | ▲ | ✓ | ▲ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P8 | <24 | m | UG | iOS | ✓✓ | ✓ | 4h01 | ✓ | ▲ | ✓ | ✓ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P9 | <24 | m | UG | iOS | ✓✓ | ✓ | 4h11 | ✓ | ▲ | ✓ | ▲ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P10 | <24 | f | M | iOS | ✓✓ | ✓ | 4h13 | ✓ | ▲ | ✓ | ✓ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P11 | ≥24 | f | PhD | Android | ✓✓ | ✓ | 4h29 | ✓ | ✓ | ✓ | ✓ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P12 | ≥24 | m | M | Android | ✓✓ | ✓ | 3h56 | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P13 | ≥24 | f | UG | iOS | ✓✓ | ✓ | 4h16 | ✓ | ▲ | ✓ | ✓ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P14 | ≥24 | f | M | Android | ✓✓ | ✓ | 4h12 | ✓ | ✓ | ✓ | ✓ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |
| P15 | ≥24 | f | PhD | iOS | ✓✓ | ✓ | 4h14 | ✓ | ▲ | ✓ | ▲ | ▲ | ▲ | △ | △ | △ | △ | △ | △ | △ | △ | △ |

people with prior educational experiences in N.America, Africa, Europe, and Asia.

Participants were each given a £50 voucher after the focus groups, before being invited to the optional follow-on.

Focus groups

We ran focus groups with three cohorts. For each cohort there were two sessions, each of approximately two-hour duration. Five students participated in each cohort, attending the first and second sessions with the same people.

In the first two-hour session, the researcher introduced the concept of digital phenotyping for mental health and then discussed collecting passive and self report data. The data covered in Table 1 was discussed, with the exception of browser history. The information was based upon sensor descriptions for the AWARE logging framework [31]. We also discussed collecting camera images along with camera interactions. The self report data discussed included the items described in Table 2 as well as demographics, medication details, and course related data.

In the second two-hour session we discussed what digital phenotyping technology might seek to infer from the data collected (e.g anxiety, stress, depression) and also what might be done with the data in terms of storage and sharing.

Tracking and interviews

The final stage of the research was optional. The participants were invited to install a tracking application onto their personal smartphone, which would record data and upload it to our database. This system was built upon the AWARE framework [31]. Based upon outcomes from the focus groups, the software allowed collection of screen, battery, app use (Android only) and location. Each student made an individual choice of which forms of data collection to allow.

We completed the study by conducting one-to-one interviews. In the interview we showed the participants visualisations of their data via a prototype app (Figure 1), and then showed ‘raw’ copies of their data in CSV format. If a student had not collected a form of data themselves, they were shown a preprepared sample of data not personal to them. In addition we showed them examples of microphone recordings, physical activity and other data (both visualised and raw). Transcripts from the focus groups were also shown and emergent themes from the analysis discussed.

Table 4 shows what data the participants shared, saw and discussed. P3 and P5 chose not to participate in the final stage (one was unresponsive to the invitation and the other “too busy”). The logging application failed on P12’s device so no data was collected from him, but he attended the interview.

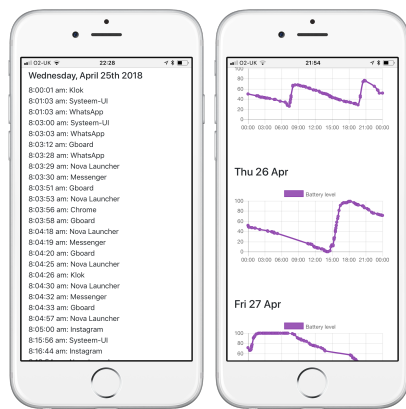


Figure 1: App prototype showing app launches and battery levels by day. Note: visualisations available to participants only during interview.

The other 12 participants installed the app for between 1 and 7 days, each generating on average 9615 rows of data.

4 FINDINGS

We transcribed the interviews and focus groups and have used thematic analysis [14] to build up an understanding of the data. Initial themes from the focus groups were discussed in the interviews and have been refined for this paper. The body of this findings section represents the key inductive themes that have emerged in the study. Deductive (i.e. pre-specified) themes about the sensor and self-report data also form an important aspect of this work, and these are represented by Tables 5 and 6. In order to produce generalisations for the deductive themes, we have used a technique called charting or matrix analysis [63, 76] (which is appropriate when there are differing opinions, understandings and levels of engagement among participants). The theoretical perspective underlying our analysis is one of "realism", simply meaning that we take the participant's opinions at face value (as opposed to looking for underlying motives or social constructs) [14]. This is appropriate for studying acceptability where subjective opinions are of importance, even if these are mistaken or underdeveloped. This perspective acknowledges that an aspect of making interventions more acceptable may be to educate and explain.

Potential for value

Our first theme concerns general opinions on the potential value of digital phenotyping technology. Most participants saw some value in the technology and all saw the need for universities and students to address mental health. P1 explained that the focus at university is on physical health, when mental health can be a bigger issue for students:

P1: "[Physical health] doesn't limit you as much in your University career as much as mental health can do."

The participants all recognised the seriousness of the issues, and had the sense that mental health and wellbeing challenges are widespread. Several discussed personal experiences and most knew at least one person who had faced problems. P5 explained that some forms of mental wellbeing issues are easy to talk to people about at university, but many are still taboo:

P5: "Anxiety and stress are much more prevalent themes in a university so it's less taboo, but if someone were to point blank ask you if you were depressed, you know, then that would be a lot harder to admit to."

However, a key issue participants raised throughout the study was not so much the difficulties of talking to others about mental health, but difficulties of recognising signs and symptoms in the first place, and then knowing what services or resources to turn to:

P2: "You don't really realise until it's really bad and then its, oh! Well what do I do now?"

Based upon this issue, participants thought that digital phenotyping technology may best help with reflection on and awareness of one's own mental health. They thought an important use for the app would be giving information back to the user and signposting to services.

P5: "An app like this ... it could at least point you in the right direction if it picks up on certain things. Cos just by having the app you would then be more aware of what you can take advantage of as a student."

The broader perspective here was that mental health is primarily a student's responsibility. In the first place, students should be supported in recognising problems themselves and deciding if and where to seek help:

P14: "I think it's better for students to give them a chance to improve themselves and then if that didn't work then maybe they can reach out to someone else, professional help in or outside of the university."

Beyond this, some participants (particularly the postgraduates in group 3), saw the value of such a technology in terms of improving research and as a tool for measuring the scope and scale of the problems faced at university to inform policy and services:

P12: "I think the technology does provide an excellent foundation for furthering other research and arguments for better resources for people ... if you manage to design it in a way that the data collection is effective and discreet, and doesn't provide more problems to people or hurt them then it would be fantastic."

Most in group 1 on the other hand did not agree that monitoring technology would be appropriate for informing

Table 5: Overview of the participants' opinions on passive data collection

| Sensor | Opinion summary | Example quotes |
|--------------------|--|--|
| Accel. / Activity | Some problems understanding what this data is and encompasses. Participants thought walking was relevant to mental health. | P12: "I think at least one of the signs of depression is lethargy and apathy, so if you go from walking your necessary 10,000 steps a day to, like, 2,000 ... it could be a sign that something's wrong." |
| App usage | Concerns about tracking specific apps, such as dating and LGBT+ related apps, but others less so. Also concerns about misinterpreting uses of some apps. Communication apps seen as relevant. Interests shown in personal tracking of app use. | P11: "a student at the College of Art ... might spend hours on Instagram because they document and they promote themselves." P5: "Tinder people might be 'there's no way I'm actually going to let you keep track when I open or close that', but I don't see a problem if you track when I play my games or when I'm on Facebook." |
| Battery / charging | Participants could not see the purpose of this for mental health, but did not see much privacy concern. | P11: "I don't care if you know about my battery." P14: "I wasn't quite sure what that has to do with the mental status." |
| Bluetooth | Main concerns raised are for other people's privacy. Students do not think they use Bluetooth often or connect to many devices. | P1 "I might consent for this but if I have my Bluetooth on and one of my friends has their Bluetooth on and they do not consent to this app, you will still get data from her phone." |
| Call logs | Main objection is on the grounds that few of the students use the telephone. Phone mainly used where there is poor data or for calling companies and services. | P9 "This caused a lot of outcry a few years back, When they realised phone companies do this ... And here we are, we've become quite tolerant about it." P7: "Most of [my friends] are in WhatsApp and Skype, I don't know. I get contacted from all kinds of apps, but not calling." |
| Camera | Any collection of content seen as unacceptable. No one would accept automatic collection of pictures, and most would not accept manual uploads (e.g. of selfie). Overuse of camera may be problematic. | P12: "You'd end up putting a piece of duct tape over both cameras." P8: "I think if it asked you to take a photo. I don't know what you could learn from that." P13: "I think it's a very unhealthy behaviour to encourage." |
| Screen | Seen as one of the less invasive ways of seeing interaction patterns and daily routines. Worries about false positives (e.g. screen on due to movement or notification). | P12 "interesting to see if there are correlations between how much screen-time you get versus your ability to sleep and your ability to relax and put it away, because the screen being on and the screen being off, it doesn't bother me for privacy issues." |
| Keyboard | Highly unacceptable to record keys clicked. Participants do not want their messages or searches tracked. Keyboard events (key press counts etc.) rather than content more acceptable. | P1 "That's a scary one... I definitely don't want them to see what I Google." |
| Location | Mixed opinions: highly unacceptable for some, but others would be happy if there is a need. Limiting tracking to campus seen as more acceptable. | P5 "It's something that drains battery and takes up ... space." P6 "Unless they show my professors how long I spend in the library I'm fine with that." P14 "if the app will work without me switching location on then I would choose not to." |
| Light sensor | Seen as relatively acceptable. Light seen as relevant to mental health. | P2 "I was really curious about it, like how does it, do that?" P12 "This one appeals to me ... the amount of exposure that people have to light can affect your mental health." |
| Mic. | Recordings very unacceptable. Quantification somewhat acceptable. | P10 "I have to ask for permission all the times whenever I like talk with someone." P14: "it's like an invasion of privacy." |
| SMS / Msg. | Message counts not generally seen as a problem, but tracking content unacceptable. Participants rarely use SMS. | P12 "I would be okay with a log as long as it didn't have any of the content." |

research or decision making, arguing that they themselves and others would not use it unless there was a direct personal benefit. A dissenting voice about the potential value of digital phenotyping was P13, who viewed most mental health technology as a "cop out", a cheap and inferior alternative to counselling and contact.

Potential for harm

The reference by P12 to "hurt" in the previous quote is important because even though all but one of the participants were positive about technology for mental health, they were also dubious that monitoring via passive sensing was the correct approach. One of the concerns underlying this was that digital phenotyping would not necessarily mean better support. In the words of P1:

P1: *"An app itself is not going to fix it."*

The point here is partly that there needs to be some sort of service behind the app; it is not enough just to make assessments, but these would need to be acted on. However, this notion of remote assessment by members of an institution worried people. For example in the words of P13:

P13: *"I feel very much like I have to protect myself and only tell them what I feel is necessary to get the support that I need. And even then I feel like it's questioned or dismissed or they just don't have the policies and procedures to properly support me."*

This is not just an issue at the study site (which the participants felt was more focused on mental health than many others). Several participants spoke about experiences elsewhere, for example:

P10: *"I was in China during my undergraduate ... Our university wanted to know our mental situation, and if the test result is bad, like err some teacher will get a task, this teacher will be responsible for this student's mental health ... So some student really don't like that, so when they fill this form, they will like err do not fill the real situation."*

Self determination was important, otherwise the system would be "infantilising" (a word used a lot in groups 1 and 2). The participants referred to going to university as a time to become adults. This was important for some because it meant gaining control:

P6: *"[At school they would] just like sort of blame you for it and just like sort of check your body for scars."*

The arguments made about mental health care as a personal responsibility, therefore, should not be seen as a moral individualism among the students, but as a response to problematic systems of care. These are systems that students want to artfully navigate or simply protect themselves from.

Another worry was not so much loss of autonomy, but the potential for discrimination based upon labels:

P11: *"This app would then give them information or data to make inferences about me that could potentially...discriminate against me as a student perhaps or label me a certain way which I'm not comfortable with."*

Strong arguments were built over the sessions about the relevance and problematics of the institutional contexts into which the technology would enter. It became clear that digital phenotyping for student mental health should not be narrowly construed as a technical or computer science problem.

Privacy concerns

As acknowledged in prior digital phenotyping research, there are privacy issues at play. However, these are not simply issues of data security.

Generally, the participants felt that many of the suggested forms of data collection were "invasive" or at least "sensitive":

P9: *"This is very sensitive data collection."*

Partly the issue was that the data may leak out from the university. Some students felt they trusted their university with their data, but others worried it would be vulnerable to hacking.

The core privacy concerns, however, were whether university staff that knew or taught the student would have access to this data.

P7: *"I wouldn't [want] my tutors to know, because if I have an issue which is affecting my work, I'll tell my tutors. I'll email them and tell them. They don't need to know everything that's going on in my life."*

One of the participants was worried lecturers might find out that she sets her alarm for 11am. It was not that the students never wanted lecturers or tutors to have access to information, but that release of information should be controlled. One idea that came up in all groups was that data could be released when needing to provide evidence for reasons of absence, or for needing deadline extensions.

Relevancy of data collected

The students felt that there must be good reasons in order for data to be collected:

P1: *"It's like I don't want you to have data that you don't absolutely need."*

Importantly, most participants did not understand why a mental health app would need to collect information that is not 'logically' related to mental health. Although it was accepted that data can be used to make inferences, just how seemingly innocuous things could be linked to mental health was not well understood:

P12: *"AI totally baffles me now."*

Table 6: Overview of the participants' opinions on self report

| Survey | Opinion summary | Example quotes |
|--------|---|--|
| PHQ9 | The questions might cause worry or bad thoughts, particularly the ones about suicide and self harm. Students were concerned what would happen (or not happen) if you say you are depressed. | P5: "There's a lot of questions, and at the end you're convinced you're depressed," P8: "they're not even thinking about it and they read this question and ... ok maybe I'm better off dead." |
| GAD7 | The questions might cause worry. Participants thought it was too easy to self mis-diagnose. Concerns about the consequences of answering these questions. | P6: "What's going to happen if I answer that I have been feeling super anxious? Is there going to be a team of medics rushed into my room?" |
| WEMWBS | Positively framed questions are more appropriate. However, answering negatively may be upsetting. | P9: "The first question is 'I've been feeling optimistic about the future', and for me it's like well uh, oh the future, I'm graduating, where will I be? Can I stay here? I want to be in the UK but I don't know." |

P11: *"I guess I don't really see the relevance, when I think in terms of a lot of data that's being collected."*

Being able to see relevance (rightly or wrongly) of data to mental health meant that the participants saw the data collection more positively (e.g. the comments about walking or light in Table 5). Even the relatively unacceptable forms of data collection may become acceptable if there is a perceived need, e.g. tracking "trigger words" in text messages or searches would be more acceptable than tracking all words.

Several participants wondered if inferences could be made on the device and then shared, rather than the raw data.

P8: *"inferences ... that's the data you can have."*

If on-device analytics is possible and participants are able to control and share the inferences, then it seems much more likely that students would accept this.

Making users worry

Regarding the self-report questions, the participants thought being asked about anxiety, depression and wellbeing may make people worry unnecessarily, or may even cause negative thoughts and feelings. The questions about anxiety and depression were "dark". The positively framed questions in the WEBWBS questionnaire were more acceptable to participants, but even positive questions may cause upset:

P13: *"if you were feeling helpless or hopeless and you answered no to ['I've been feeling useful'] I think it would just make your symptoms worse."*

A major criticism of the surveys, particularly the depression and anxiety surveys, was that these would be better used in a face-to-face setting with a trained person.

P9: *"These questions may be raised by a medical professional more appropriately than through an app."*

The participants made suggestions that more abstract mood tracking may be preferable to having to answer these questions, or something such as bullet journaling.

Effect on device

A key concern about passive data collection was the effect this would have on their device. A fundamental concern for the participants was their device's battery. They argued any app that negatively effects battery life is likely not to be used unless it is of high value. Use of data and storage were also concerns, although less pronounced than battery. The participants all regularly connected to University Wi-Fi, and for some this was their only source of data.

Issues of whether it is right to expect students to have a smartphone and for this to then be used for a university initiative were also questioned. Potentially it would be better to supply students with new smartphones.

5 DISCUSSION

This study raises doubts about the acceptability to students of digital phenotyping by universities. However, acceptability is not a simple yes/no question and in this section we will discuss these doubts and identify ways in which they may be overcome.

Acceptability of digital phenotyping

As described earlier, the Theoretical Framework for Acceptability (TFA) [71] is appropriate for structuring our findings with respect to acceptability. The TFA has seven components, which we will step through here:

Affective attitude. This concerns how people feel about and may be affected by digital phenotyping technology. On the positive side we found that all participants thought that universities and students should take mental health seriously, and that most were in favour of technology-based approaches

such as apps. However, participants pointed out that the proposed technology may cause people worry and anxiety by the questions asked and in terms of what lecturers and tutors may find out. They were also concerned about a loss of autonomy and dignity in navigating systems of care.

Developers of digital phenotyping systems need to be transparent and careful about what data is collected and who has access, and need to consider the affective user experience of having and using the technology.

Burden. This concerns the perceived amount of effort required for using the technology. There was no suggestion among participants that having an app such as this would place specific burden on them, although several did question whether they would make the effort to install an app and answer questions. There were also worries about needing to charge batteries frequently.

Developers of digital phenotyping technology should be aware that low burden characterises digital phenotyping, but this alone will not ensure acceptability.

Ethicality. This is the extent to which digital phenotyping fits with individuals' value systems. The students' key concerns were loss of autonomy, control and dignity. The transition to adulthood was an important aspect of many of the younger students' considerations.

Developers should provide controls over how information is released, and on-device analytics with selective sharing of inferences rather than raw data is potentially fruitful here.

Coherence. This concerns whether people understand the intervention and how it works. Many participants were aware that inferences could be made about them from behavioural data, but wanted to see logical relationships between a datatype and wellbeing. They were not aware of how much could be extrapolated from seemingly innocuous data such as battery charge over time.

Developers need to carefully explain why data is collected and how it holds meanings. Case-based examples might help to ensure that consent is informed as well as address any mismatch between perceived and actual potential threats. Consent that is oriented to inferences rather than just the types of raw data may also be needed.

Opportunity costs. This concerns what is given up to exchange in the intervention. In the case of smartphone-based digital phenotyping, the fundamental concern to students is the effect on their battery. Data, storage and performance costs are also concerns.

Developers of digital phenotyping systems should minimise the effects on participants' devices, particularly battery. Although smartphone penetration is high, it should not be assumed students own new or high-end devices, or can charge

batteries regularly. It may be reasonable to supply students with new devices to use.

Perceived effectiveness. This concerns whether the intervention is perceived by participants as likely to achieve its purpose. The overriding factor here is not whether participants think digital phenotyping will make correct assessments, but how they see it situated within a system of care. The students pointed out that these systems of care need to be navigated in order that students get the help they think they need and to avoid what they do not want.

Designers of digital phenotyping technology need to engage in service design rather than just technology design. Technologies should be appropriate for the artful navigation of care, and/or be part of a reconfiguration of existing institutional approaches.

Self efficacy. This concerns whether users are confident they can make changes. The participants did not report feeling in control of their mental health. They had difficulties recognising when things were wrong or knowing what to do.

Designers of digital phenotyping technology should provide information about mental health and wellbeing and support reflection for self awareness. Supporting know-how for change and expediting human contact may be valuable.

Future work

Design work in this area should be more user centered; the design recommendations we have outlined above could inform such work. Further qualitative work could involve larger sample sizes and/or could focus on: subpopulations who face additional mental health challenges at university e.g. LGBT+; people with diagnosed disorders; or, surveillance and sharing within clinical, peer, and family contexts.

6 CONCLUSION

We have looked at digital phenotyping through the lens of *acceptability* in order to develop a sense of what engenders willing participation by students in data collection. We have uncovered a range of views and beliefs that suggest seeing digital phenotyping not as a technical or computer science challenge of data collection and analytics, but as an interdisciplinary design challenge in which the ways in which students are supported are rethought. There are important technical challenges still to address, but if we are not paying attention at the same time to the contexts of care for mental health and wellbeing, and if we are not putting student autonomy and self determination at the heart of our approaches, then innovations in this area may be in vain.

7 ACKNOWLEDGEMENT

We received funding for this work from the University of Edinburgh Challenge Investment fund.

REFERENCES

- [1] Saeed Abdullah, Mark Matthews, Elizabeth L. Murnane, Geri Gay, and Tanzeem Choudhury. 2014. Towards Circadian Computing: "Early to Bed and Early to Rise" Makes Some of Us Unhealthy and Sleep Deprived. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 673–684.
- [2] Talayah Aledavood, Ana Maria Triana Hoyos, Tuomas Alakörkkö, Kimmo Kaski, Jari Saramäki, Erkki Isometsä, and Richard K Darst. 2017. Data collection for mental health studies through digital platforms: requirements and design of a prototype. *JMIR research protocols* 6, 6 (2017).
- [3] David Armstrong. 1995. The rise of surveillance medicine. *Sociology of health & illness* 17, 3 (1995), 393–404.
- [4] Joost Asselbergs, Jeroen Ruwaard, Michal Ejdy, Niels Schrader, Marit Sijbrandij, and Heleen Riper. 2016. Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study. *Journal of medical Internet research* 18, 3 (2016).
- [5] Shrey Bagroy, Ponnurangam Kumaraguru, and Munmun De Choudhury. 2017. A Social Media Based Index of Mental Well-Being in College Campuses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1634–1646. <https://doi.org/10.1145/3025453.3025909>
- [6] Lyndsey L. Bakewell, Konstantina Vasileiou, Kiel S. Long, Mark Atkinson, Helen Rice, Manuela Barreto, Julie Barnett, Michael Wilson, Shaun Lawson, and John Vines. 2018. Everything We Do, Everything We Press: Data-Driven Remote Performance Management in a Mobile Workplace. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 371, 14 pages. <https://doi.org/10.1145/3173574.3173945>
- [7] Ian Barnett, John Torous, Patrick Staples, Luis Sandoval, Matcheri Keshavan, and Jukka-Pekka Onnela. 2018. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. *Neuropsychopharmacology* (2018), 1.
- [8] Dennis Becker, Vincent Bremer, Burkhardt Funk, Joost Asselbergs, Heleen Riper, and Jeroen Ruwaard. 2016. How to Predict Mood? Delving into Features of Smartphone-Based Data. *Twenty-second Americas Conference on Information Systems* (2016), 1–10.
- [9] Dror Ben-Zeev, Emily A Scherer, Rui Wang, Haiyi Xie, and Andrew T Campbell. 2015. Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatric rehabilitation journal* 38, 3 (2015), 218.
- [10] Dror Ben-Zeev, Rui Wang, Saeed Abdullah, Rachel Brian, Emily A Scherer, Lisa A Mistler, Marta Hauser, John M Kane, Andrew Campbell, and Tanzeem Choudhury. 2015. Mobile behavioral sensing for outpatients and inpatients with schizophrenia. *Psychiatric services* 67, 5 (2015), 558–561.
- [11] Matthew Billingsley. 2015. More than 80% of medical students with mental health issues feel under-supported, says Student BMJ survey. *Student BMJ* 23 (2015), h4521.
- [12] Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. 2011. Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 47–56. <https://doi.org/10.1145/2037373.2037383>
- [13] Mehdi Boukhechba, Yu Huang, Philip Chow, Karl Fua, Bethany A. Teachman, and Laura E. Barnes. 2017. Monitoring Social Anxiety from Mobility and Communication Patterns. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17)*. ACM, New York, NY, USA, 749–753.
- [14] Virginia Braun, Victoria Clarke, and Gareth Terry. 2014. Thematic analysis. *Qual Res Clin Health Psychol* 24 (2014), 95–114.
- [15] Emma Broglia, Abigail Millings, and Michael Barkham. 2017. Challenges to addressing student mental health in embedded counselling services: a survey of UK higher and further education institutions. *British Journal of Guidance & Counselling* 0, 0 (2017), 1–15. <https://doi.org/10.1080/03069885.2017.1370695> arXiv:<https://doi.org/10.1080/03069885.2017.1370695>
- [16] Poppy Brown. 2016. *The Invisible Problem?: Improving Students' Mental Health*. Higher Education Policy Institute.
- [17] Michelle Nicole Burns, Mark Begale, Jennifer Duffecy, Darren Gergle, Chris J Karr, Emily Giangrande, and David C Mohr. 2011. Harnessing context sensing to develop a mobile intervention for depression. *Journal of medical Internet research* 13, 3 (2011).
- [18] Luca Canzian and Mirco Musolesi. 2015. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 1293–1304.
- [19] Laura L Carstensen, Bulent Turan, Susanne Scheibe, Nilam Ram, Hal Ersner-Hersfield, Gregory R Samanez-Larkin, Kathryn P Brooks, and John R Nesselroade. 2011. Emotional experience improves with age: evidence based on over 10 years of experience sampling. *Psychology and aging* 26, 1 (2011), 21.
- [20] Larry Chan, Vedant Das Swain, Christina Kelley, Kaya de Barbaro, Gregory D. Abowd, and Lauren Wilcox. 2018. Students' Experiences with Ecological Momentary Assessment Tools to Report on Emotional Well-being. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1 (March 2018), 3:1–3:20.
- [21] Zhenyu Chen, Mu Lin, Fanglin Chen, Nicholas D. Lane, Giuseppe Cardone, Rui Wang, Tianxing Li, Yiqiang Chen, Tanzeem Choudhury, and Andrew T. Campbell. 2013. Unobtrusive Sleep Monitoring Using Smartphones. In *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '13)*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, 145–152.
- [22] Chia-Fang Chung, Nanna Gorm, Irina A. Shklovski, and Sean Munson. 2017. Finding the Right Fit: Understanding Health Tracking in Workplace Wellness Programs. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4875–4886. <https://doi.org/10.1145/3025453.3025510>
- [23] Karen Church, Denzil Ferreira, Nikola Banovic, and Kent Lyons. 2015. Understanding the Challenges of Mobile Phone Usage Data. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 504–514. <https://doi.org/10.1145/2785830.2785891>
- [24] Richard Dawkins. 1982. *The Extended Phenotype*. Oxford University Press.
- [25] Kevin Eagan, Ellen Bara Stolzenberg, Hilary B Zimmerman, Melissa C Aragon, Hannah Whang Sayson, and Cecilia Rios-Aguilar. 2017. The American Freshman: National Norms Fall 2016. Retrieved August 12, 2018 from https://www.insidehighered.com/sites/default/server_files/files/TheAmericanFreshman2016PREPUB.pdf
- [26] Nathan Eagle and Alex Pentland. 2006. Reality mining: sensing complex social systems. *Personal and ubiquitous computing* 10, 4 (2006), 255–268.
- [27] Paul Eskes, Marco Spruit, Sjaak Brinkkemper, Jacob Vorstman, and Martien J Kas. 2016. The sociability score: App-based social profiling from a healthcare perspective. *Computers in Human Behavior* 59 (2016), 39–48.

- [28] Asma Ahmad Farhan, Chaoqun Yue, Reynaldo Morillo, Shweta Ware, Jin Lu, Jinbo Bi, Jayesh Kamath, Alexander Russell, Athanasios Bamis, and Bing Wang. 2016. Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data. In *2016 IEEE Wireless Health (WH)*. 1–8.
- [29] Raihana Ferdous, Venet Osmani, and Oscar Mayora. 2018. Smartphone apps usage patterns as a predictor of perceived stress levels at workplace. *arXiv preprint arXiv:1803.03863* (2018).
- [30] Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, Louise Barkhuus, and Anind K. Dey. 2014. Contextual Experience Sampling of Mobile Application Micro-usage. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 91–100. <https://doi.org/10.1145/2628363.2628367>
- [31] Denzil Ferreira, Vassilis Kostakos, and Anind K. Dey. 2015. AWARE: Mobile Context Instrumentation Framework. *Frontiers in ICT* 2 (2015), 6.
- [32] Office for National Statistics. 2018. Estimating suicide among higher education students, England and Wales: Experimental Statistics.
- [33] Christopher Frauenberger, Marjo Rauhala, and Geraldine Fitzpatrick. 2017. In-action ethics. *Interacting with Computers* 29, 2 (2017), 220–236.
- [34] Robert R German, LM Lee, J Horan, R Milstein, C Pertowski, M Waller, et al. 2001. Updated guidelines for evaluating public health surveillance systems. *MMWR Recomm Rep* 50, 1-35 (2001).
- [35] Joel M Hektner, Jennifer A Schmidt, and Mihaly Csikszentmihalyi. 2007. *Experience sampling method: Measuring the quality of everyday life*. Sage.
- [36] Yu Huang, Haoyi Xiong, Kevin Leach, Yuyan Zhang, Philip Chow, Karl Fua, Bethany A. Teachman, and Laura E. Barnes. 2016. Assessing Social Anxiety Using Gps Trajectories and Point-of-interest Data. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 898–903.
- [37] Galen Chin-Lun Hung, Pei-Ching Yang, Chia-Chi Chang, Jung-Hsien Chiang, and Ying-Yeh Chen. 2016. Predicting negative emotions based on mobile phone usage patterns: an exploratory study. *JMIR research protocols* 5, 3 (2016).
- [38] Josiah Gabriel Hunt. 2015. Suicide mortality among students in South Korea: An extended discussion. In *International Forum*, Vol. 18. 105–120.
- [39] TR Insel. 2017. Digital phenotyping: Technology for a new science of behavior. *JAMA* 318, 13 (2017), 1215–1216. <https://doi.org/10.1001/jama.2017.11295>
- [40] Sachin H Jain, Brian W Powers, Jared B Hawkins, and John S Brownstein. 2015. The digital phenotype. *Nature Biotechnology* 33, 5 (2015), 462.
- [41] Christina Kelley, Bongshin Lee, and Lauren Wilcox. 2017. Self-tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 629–641. <https://doi.org/10.1145/3025453.3025750>
- [42] Helen Kerr. 2013. NUS Mental Distress Survey Overview. Retrieved August 12, 2018 from <https://www.nus.org.uk/Global/Campaigns/20130517MentalDistressSurveyOverview.pdf>
- [43] Kurt Kroenke and Robert L Spitzer. 2002. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric annals* 32, 9 (2002), 509–515.
- [44] Kwangyoung Lee and Hwajung Hong. 2018. MindNavigator: Exploring the Stress and Self-Interventions for Mental Wellness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 572, 14 pages. <https://doi.org/10.1145/3173574.3174146>
- [45] Uichin Lee, Joonwon Lee, Minsam Ko, Changhun Lee, Yuhwan Kim, Subin Yang, Koji Yatani, Gahgene Gweon, Kyong-Mee Chung, and June-hwa Song. 2014. Hooked on smartphones: an exploratory study on smartphone overuse among college students. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2327–2336.
- [46] Robert LiKamWa, Yunxin Liu, Nicholas D Lane, and Lin Zhong. 2013. Moodscope: Building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. ACM, 389–402.
- [47] Melanie Lovatt and John Holmes. 2017. Digital phenotyping and sociological perspectives in a Brave New World. *Addiction* 112, 7 (2017), 1286–1289.
- [48] Ewa Luger and Tom Rodden. 2013. An Informed View on Consent for UbiComp. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 529–538. <https://doi.org/10.1145/2493432.2493446>
- [49] Ann Macaskill. 2013. The mental health of university students in the United Kingdom. *British Journal of Guidance & Counselling* 41, 4 (2013), 426–441.
- [50] Anmol Madan, Manuel Cebrian, David Lazer, and Alex Pentland. 2010. Social sensing for epidemiological behavior change. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, 291–300.
- [51] Gloria Mark, Shamsi T. Iqbal, Mary Czerwinski, Paul Johns, and Akane Sano. 2016. Neurotics Can'T Focus: An in Situ Study of Online Multitasking in the Workplace. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1739–1744. <https://doi.org/10.1145/2858036.2858202>
- [52] Gloria Mark, Yiran Wang, and Melissa Niiya. 2014. Stress and Multitasking in Everyday College Life: An Empirical Study of Online Activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 41–50. <https://doi.org/10.1145/2556288.2557361>
- [53] Mark Matthews, Elizabeth Murnane, Jaime Snyder, Shion Guha, Pamara Chang, Gavin Doherty, and Geri Gay. 2017. The double-edged sword: A mixed methods study of the interplay between bipolar disorder and technology use. *Computers in Human Behavior* 75 (2017), 288–300.
- [54] Donald McMillan, Alistair Morrison, and Matthew Chalmers. 2013. Categorised Ethical Guidelines for Large Scale Mobile HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1853–1862. <https://doi.org/10.1145/2470654.2466245>
- [55] Abhinav Mehrotra, Fani Tsapeli, Robert Hendley, and Mirco Musolesi. 2017. MyTraces: investigating correlation and causation between users' emotional states and mobile phone interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 83.
- [56] Graham F Moore, Suzanne Audrey, Mary Barker, Lyndal Bond, Chris Bonell, Wendy Hardeman, Laurence Moore, Alicia O'Cathain, Tanaze Tinati, Daniel Wight, et al. 2015. Process evaluation of complex interventions: Medical Research Council guidance. *bmj* 350 (2015), h1258.
- [57] Alistair Morrison, Xiaoyu Xiong, Matthew Higgs, Marek Bell, and Matthew Chalmers. 2018. A Large-Scale Study of iPhone App Launch Behaviour. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 344, 13 pages. <https://doi.org/10.1145/3173574.3173918>
- [58] Ian Neale, Laura Piggott, Josephine Hansom, and Sam Fagence. 2016. Student Resilience: Unite Students Insight Report. Retrieved August 12, 2018 from <https://www.unitestudents.com/about-us/insightreport/>

- 2016-full-report
- [59] Gina Neff and Dawn Nafus. 2016. *Self-Tracking*. MIT Press.
 - [60] Alicia L. Nobles, Jeffrey J. Glenn, Kamran Kowsari, Bethany A. Teachman, and Laura E. Barnes. 2018. Identification of Imminent Suicide Risk Among Young Adults Using Text Messages. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 413, 11 pages. <https://doi.org/10.1145/3173574.3173987>
 - [61] Aisling Ann O’Kane, Yvonne Rogers, and Ann E. Blandford. 2015. Concealing or Revealing Mobile Medical Devices?: Designing for Onstage and Offstage Presentation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1689–1698. <https://doi.org/10.1145/2702123.2702453>
 - [62] Jukka-Pekka Onnela and Scott L Rauch. 2016. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* 41, 7 (2016), 1691.
 - [63] Anthony J Onwuegbuzie, Wendy B Dickinson, Nancy L Leech, and Annmarie G Zoran. 2009. A qualitative framework for collecting and analyzing data in focus group research. *International journal of qualitative methods* 8, 3 (2009), 1–21.
 - [64] Alex Pentland. 2015. *Social Physics: How social networks can make us smarter*. Penguin.
 - [65] Gary W. Pritchard, Pam Briggs, John Vines, and Patrick Olivier. 2015. How to Drive a London Bus: Measuring Performance in a Mobile and Remote Workplace. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1885–1894. <https://doi.org/10.1145/2702123.2702307>
 - [66] Debra Rickwood, Nic Telford, Shaunagh O’Sullivan, Dimity Crisp, and Robby Magyar. 2016. National Tertiary Student Wellbeing Survey 2016. Retrieved August 12, 2018 from <https://www.headspace.org.au/assets/Uploads/headspace-NUS-Publication-Digital.pdf>
 - [67] Bart Rienties and Bethany Alden Rivers. 2014. Measuring and understanding learner emotions: Evidence and prospects. *Learning Analytics Review* 1 (2014), 1–28.
 - [68] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal Tracking As Lived Informatics. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1163–1172. <https://doi.org/10.1145/2556288.2557039>
 - [69] Sohrab Saeb, Mi Zhang, Christopher J Karr, Stephen M Schueller, Marya E Corden, Konrad P Kording, and David C Mohr. 2015. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *Journal of medical Internet research* 17, 7 (2015).
 - [70] Koustuv Saha and Munmun De Choudhury. 2017. Modeling Stress with Social Media Around Incidents of Gun Violence on College Campuses. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 92 (Dec. 2017), 27 pages. <https://doi.org/10.1145/3134727>
 - [71] Mandeep Sekhon, Martin Cartwright, and Jill J Francis. 2017. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC health services research* 17, 1 (2017), 88.
 - [72] Souraya Sidani, Dana R Epstein, Richard R Bootzin, Patricia Moritz, and Joyal Miranda. 2009. Assessment of preferences for treatment: validation of a measure. *Research in nursing & health* 32, 4 (2009), 419–431.
 - [73] Vivek K Singh and Rishav R Agarwal. 2016. Cooperative phoneotypes: exploring phone-based behavioral markers of cooperation. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 646–657.
 - [74] Vivek K. Singh and Isha Ghosh. 2017. Inferring Individual Social Capital Automatically via Phone Logs. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW (Dec. 2017), 95:1–95:12.
 - [75] Andrew L Skinner, Angela S Attwood, Roland Baddeley, Karen Evans-Reeves, Linda Bauld, and Marcus R Munafo. 2017. Digital phenotyping and the development and delivery of health guidelines and behaviour change interventions. *Addiction* 112, 7 (2017), 1281–1285.
 - [76] Liz Spencer and Jane Ritchie. 2002. Qualitative data analysis for applied policy research. In *Analyzing qualitative data*. Routledge, 187–208.
 - [77] Robert L Spitzer, Kurt Kroenke, Janet BW Williams, and Bernd Löwe. 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine* 166, 10 (2006), 1092–1097.
 - [78] Thomas Stütz, Thomas Kowar, Michael Kager, Martin Tiefengraber, Markus Stuppner, Jens Blechert, Frank H Wilhelm, and Simon Ginzinger. 2015. Smartphone based stress prediction. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 240–251.
 - [79] Ruth Tennant, Louise Hiller, Ruth Fishwick, Stephen Platt, Stephen Joseph, Scott Weich, Jane Parkinson, Jenny Secker, and Sarah Stewart-Brown. 2007. The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health and Quality of life Outcomes* 5, 1 (2007), 63.
 - [80] TR Thorley. 2017. Not By Degrees: Improving Student Mental Health In The UKs Universities. Retrieved August 12, 2018 from https://www.ippr.org/files/2017-09/1504645674_not-by-degrees170905.pdf
 - [81] John Torous, Mathew V Kiang, Jeanette Lorme, and Jukka-Pekka Onnela. 2016. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. *JMIR mental health* 3, 2 (2016).
 - [82] Jean M Twenge, Brittany Gentile, C Nathan DeWall, Debbie Ma, Katharine Lacefield, and David R Schurtz. 2010. Birth cohort increases in psychopathology among young Americans, 1938–2007: A cross-temporal meta-analysis of the MMPI. *Clinical psychology review* 30, 2 (2010), 145–154.
 - [83] Universities UK. 2018. Minding Our Future: Starting a Conversation About the Support of Student Mental Health. Retrieved August 12, 2018 from <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2018/minding-our-future-starting-conversation-student-mental-health.pdf>
 - [84] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T. Campbell. 2014. StudentLife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students Using Smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 3–14. <https://doi.org/10.1145/2632048.2632054>
 - [85] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 3–14.
 - [86] Rui Wang, Weichen Wang, Min SH Aung, Dror Ben-Zeev, Rachel Brian, Andrew T Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Emily A Scherer, et al. 2017. Predicting Symptom Trajectories of Schizophrenia using Mobile Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 110.
 - [87] Rui Wang, Weichen Wang, Min S. H. Aung, Dror Ben-Zeev, Rachel Brian, Andrew T. Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Emily A. Scherer, and Megan Walsh. 2017. Predicting Symptom Trajectories of Schizophrenia Using Mobile Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 110 (Sept. 2017),

- 24 pages. <https://doi.org/10.1145/3130976>
- [88] Pascal Welke, Ionut Andone, Konrad Blaszkiewicz, and Alexander Markowetz. 2016. Differentiating Smartphone Users by App Usage. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 519–523. <https://doi.org/10.1145/2971648.2971707>
- [89] Julie Rico Williamson. 2011. Send Me Bubbles: Multimodal Performance and Social Acceptability. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 899–904. <https://doi.org/10.1145/1979742.1979513>
- [90] Martina Ziefle and Carsten Röcker. 2010. Acceptance of pervasive healthcare systems: A comparison of different implementation concepts. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference*. IEEE, 1–6.