Exploring Temporal Behaviour of App Users Completing Ecological Momentary Assessments using Mental Health Scales and Mood Logs

Raymond Bond[a] , Anne Moorhead[b], Maurice Mulvenna[a], Siobhan O'Neill[c], Courtney Potts[a], Nuala Murphy[d]

[a]Ulster University, School of Computing, Newtownabbey, UK; [b]Ulster University, School of Communication and Media, Newtownabbey, UK; [c]Ulster University, School of Psychology, Coleraine, UK; [d]Happy Mummy, Happy Baby, Belfast, UK

ABSTRACT
Smartphone-based digital phenotyping can provide insight into mood, cognition and behaviour. In this study, data analytics was carried out with data generated from a maternal mental health app to address the following question: what is the temporal behaviour of users when completing ecological momentary assessments (EMAs) with EMAs in the form of mental health scales versus EMAs in the form of mood logs? The methodology involved using the Health Interaction Log Data Analytics (HILDA) pipeline to analyse 1,461 app users. Clustering was used to characterize archetypical user engagement with the two forms of EMA. Users preferred mood log EMAs, with 6,993 mood log completions compared to 2,129 scale completions. Users are more willing to log moods at 9am and 12pm and complete mental health scales between 8pm and 10pm. The fewest number of mood logs and scale completions take place on Saturday followed by a Sunday. Whilst 'happiness' is the dominant mood during day times, 'anxiety' and 'sadness' peak during night times. The overall findings are that users prefer completing mood log EMAs and that the temporal behaviour of users engaging with EMAs in the form of mental health scales are distinctly different from how they engage with mood logs.

word count: 4,485

1. Introduction

Maternal mental health is an important and increasing problem. However, there are limited resources to monitor and provide support to mothers and their families in the area of maternal mental health. Pregnancy and the arrival of a newborn is a very exciting time in any family, it is also a time when emotions, both positive and negative, may impact on the health and well-being of the mother and family. According to the World Health Organisation (*1*), worldwide approximately 10% of pregnant women and 13% of women who have just given birth experience a mental health problem, primarily depression. Within the UK, up to one in five women develop mental health problems during pregnancy or in the first year after childbirth (*2*). The Royal College of Obstetricians and Gynaecologists' survey of 2,300 women revealed that 81% of

---

CONTACT R. Bond. Email: rb.bond@ulster.ac.uk

women had experienced a maternal mental health problem (2). The pain this causes women and their families has a negative impact on their health and well-being and a financial cost (3), as it is estimated that maternal mental health problems cost the UK 8.1 billion pounds each year (4).

Emotions are strong feelings deriving from one's circumstances, mood, or relationships with others. The literature clearly states that it is normal for expectant mothers to have emotional changes due to hormones during pregnancy. In addition, emotions during pregnancy can impact upon decision making of the individual (5). The six most reported types of emotions during pregnancy and birth are: surprise, joy, anger, fear, love, and sadness (5–7). From these six emotions, half are negative and half are positive, thus indicating that pregnancy and birth can produce mixed feelings. However, the literature emphasises more on negative than positive emotions. This is particularly true for the emotions of mothers after giving birth. The negative emotions result in well reported baby blues, postnatal depression and the most severe, postpartum psychosis (8). It is clear that emotions among mothers need to be monitored in order to improve maternal mental health.

The Royal College of Midwives (5) surveyed midwives and maternity support workers who reported that emotional support should be the main focus of postnatal care (midwives 61% and maternity support workers 78%). These figures were significantly higher than clinical observation (midwives 20% and maternity support workers 13%) and health promotion (midwives 19% and maternity support workers 8%). In total, 41% of midwives and 40% of maternity support workers reported that there is usually enough time and resources to support, and inform women on maternal emotional well-being. These results clearly indicate that there is not enough time and resources for maternal emotional well-being but it should be the main focus of postnatal care. Thus, it is a clear message from the Royal College of Midwives that emotional support should be the main focus of postnatal care and there needs to be more staff time and resources employed in this area. This survey is further supported by more recent findings from the Maternal Mental Health Alliance (2) that in almost half of the UK, pregnant women and new mothers have no access to specialist community maternal mental health services.

NICE Guidelines (9) in the UK on antenatal and postnatal mental health quality standard (QS115) states that *"women are asked about their emotional well-being at each routine antenatal and postnatal contact"*. The guideline states the quality measures, which includes evidence of arrangements for healthcare professionals to ask women about their emotional well-being at all routine antenatal and postnatal contacts. The stated outcomes include women's satisfaction with being able to discuss any concerns or worries at routine appointments, and identification of those living with mental health problems. Although this guideline is clear, there are questions around the implementations and the available resources for its effectiveness and success.

The literature clearly states that maternal mental health is an increasing problem but there is a lack of resources. Thus there is a need for more accessible resources for mothers and their families to enhance their mental health.

2. Digital phenotyping

Digital health products and services typically include mobile device apps as well as browser-based apps to a lesser extent, and can include telephony-based services, text-based chatbots and voice activated chatbots. Many of these digital well-being products

and services are simultaneously available across many channels in order to maximise availability for users.

Digital interaction technologies offer useful methods for real time data capture of the interactions of users with the products and services. We have become accustomed to using tools such as Google Analytics in order to generate aggregated usage reporting for our websites, but this kind of log analysis is available for all the digital interaction technologies we design. Indeed, we have the ability to design what data are recorded, how and where it may be stored, and crucially, how it can be analysed to reveal individual or collective usage patterns.

The main focus of previous research on the analysis of usage logs for digital interaction technologies used in health and well-being has been to aid in usability analysis (10) or to reveal usage patterns in using technology (11). Research has also been carried out to explore how rehabilitation devices can have data or event logging incorporated, but this has been more to support the goal of device monitoring (12). More recent research has examined engagement data in web-based intervention platforms but has primarily focused on visualisation of the log data (13).

Nowadays, digital phenotyping can harness the power of various devices to enable clinicians to make better decisions about treatment and care. Digital phenotyping is the term used to describe methods of quantifying individual-level human behaviour *in situ* from personal digital devices in real-time (14). This data is typically collected using smartphones which capture various forms of active data that require input from the user, for instance surveys, and passive data which is automatically generated such as GPS tracking (15). Acquiring data is the initial step in digital phenotyping, followed by data analysis to gain insights about the user (16). Early work has shown the use of digital phenotyping in mental health care is promising (17) and it is clear smartphone-based digital phenotyping has the potential to transform global mental health. Identifying individuals at risk or in need of treatment can be difficult as, for example, a new mother experiencing depression may look different when feeding during the night compared to what she reports to her clinician the following day (18). Due to the nature of mental illness it is often difficult to seek help and those that do report to a clinican often arrive after some delay (18). Therefore, digital phenotyping could be used to address the challenges in the field by predicting or determining if a person has a mental health problem.

The use of digital solutions in mental health care has been recognised as a valuable resource. Many national health departments including the UK's National Health Service (NHS) are looking to use digital well-being technologies such as health apps for self-management of diseases. Thus, logging user interactions would allow for greater insight into user needs and may provide ideas for improving these digital interventions, for example through enhanced personalisation. The NHS would benefit since the data can be automatically and hence cost-effectively collected. Such data may facilitate new ways for epidemiological analyses and provide data to inform health policies. If the NHS promote health apps and log analysis is insightful, then perhaps there is a need for a standard to maximise the utility of recorded event logs for analysis in healthcare contexts.

However, the ethical issues in using digital well-being data are being increasingly recognised and the protection of personal data is a significant and real concern. While the European Union's General Data Protection Regulation 2016/679 (GDPR) offers clear guidance including support for the principle of data minimisation, where "organisations must only process the personal data that they need to achieve its processing purposes". Using behavioural log data for research purposes would likely require dero-

gation of the regulation under GDPR Article 89 for scientific research purposes, the explicit consent by users who have been informed about the use of the data for analysis for research purposes, or both.

2.1. *Ecological Momentary Assessment*

The Experience Sampling Method (ESM) or the Ecological Momentary Assessment (EMA) is a subset of digital phenotyping that can be used to gain insight about the user. EMA originally made use of paper-diary techniques to enable people to record their observations or answers to specific questions and combined the ecological validity with the rigorous measurement techniques of psychometric research (*19*). EMA secures data about both behavioural and intrapsychic aspects of individuals' daily activities, and it obtains reports about the experience as it occurs, thereby minimizing the effects of reliance on memory and reconstruction (*20*, *21*). Nowadays, smartphones enable the capture of implicit usage data from integrated sensors such as, for example, accelerometers, and these life logging and quantified-self techniques have gained increasing popularity as smartphones become ubiquitous in daily life (*22*). Research has successfully shown how the use of these techniques combined with EMA can validate their integrated use, including works by Intille et al. (*23*), the SocioXensor system (*24*) and the iHABIT platform (*25*). EMAs and ecological momentary interventions (EMIs) have been proposed to assess and deliver 24-7 mental health support to patients (*26*). In addition, previous work has looked at the development of self-managed EMA (*27*) along with tools which allow the individual to analyse and visualise behavioural patterns to establish their own EMIs (*28*). EMA can also be used to gain insight into the user experience of a digital product and subsequent cognitive ergonomic design.

In this study, data analytics was carried out on maternal mental health data generated from the Moment Health app. This study addresses the following question: *What is the temporal behaviour of users when completing EMAs in the form of scales compared to EMAs in the form of mood logs using a maternal mental health app?*

3. Methods

This research study used a data analytical approach to study the cognitive ergonomics and user engagement of two forms of EMA - one EMA model consisting of mental health scales (up to 10 questions in one sitting), and the other EMA in the form of mood logs.

Figure 1 shows three screenshots of the Moment Health app. The screenshots show the first question of the Edinburgh Postnatal Depression Scale (EPDS) scale (*29*), the mood log screen and the quantified temporal visualisation of mood logs for the purpose of user feedback.

3.1. *Source of data*

The source of data is from the Moment Health app, which has been developed with expertise from clinicians and healthcare professionals and is aimed at new and expectant mothers and their families. It screens for prenatal and postnatal depression and associated anxieties and includes additional features such as a helpful guide to practical and accessible coping strategies. When a parent or parent-to-be downloads the
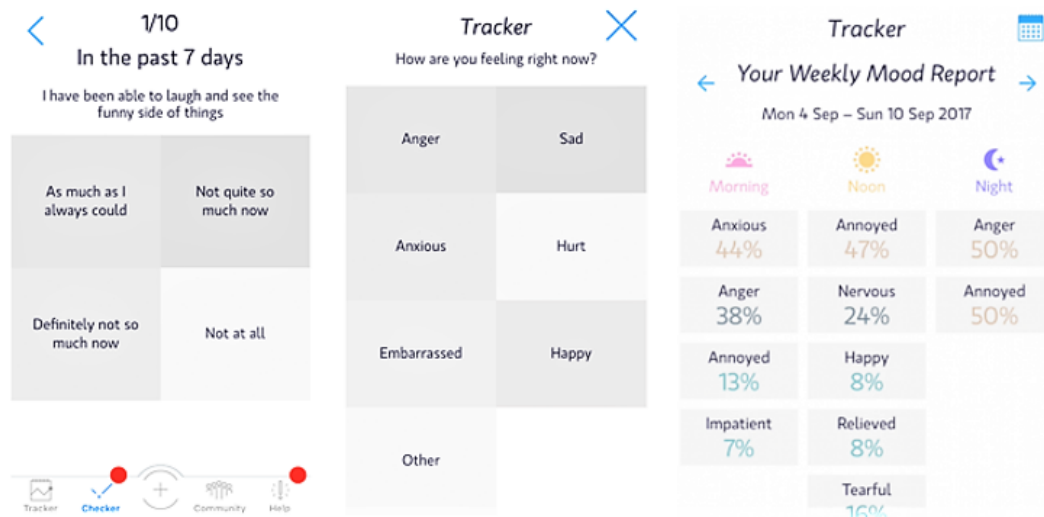
Figure 1. Screenshots of the Moment Health app mental health scales and mood logs.

app, they have access to several key, easy-to-use tools including an early intervention symptoms checker, a mood tracker, and a location tool so they can find the resources they need that are close to them, and more. This app is available for download via the Apple Store and Google play. The app consists of four key features: 1. tracker, 2. checker, 3. locater and, 4. an online community. The tracker feature includes EMAs in the form of validated mental health scales that are frequently administered in clinical practice, e.g. EPDS (*29*), GAD-7 (*30*), PHQ-9 (*31*).

The Edinburgh Postnatal Depression Scale (EPDS) is a 10-item questionnaire that was developed to identify women who have postpartum depression (*29*). The items of the scale correspond to various clinical depression symptoms, such as guilt feeling, sleep disturbance, low energy, anhedonia and suicidal ideation. Overall assessment is by total score, which is determined by adding together the scores for each of the 10 items. Higher scores indicate more depressive symptoms. The scale may be used within 8 weeks postpartum and it also can be applied for depression screening during pregnancy. It is a widely used depression screening tool, which has been adapted and validated in many languages.

The Generalized Anxiety Disorder 7 (GAD-7) scale is a 7-item self-reported questionnaire for screening and measurement of generalized anxiety disorder (*30*). GAD-7 assessment is indicated by the total score, which made up by adding together the scores for the scale all seven items.

The Patient Health Questionnaire (PHQ) is a self-administered version of the PRIME-MD diagnostic instrument for identifying those living with common mental health problems (*31*). PHQ-9 is the depression module, which scores each of the 9 DSM-IV criteria as "0" (not at all) to "3" (nearly every day). PHQ-9 is seen as a reliable and valid measure of depression severity. These characteristics plus its brevity make the PHQ-9 a useful clinical and research tool. It is not a screening tool for depression but it is used to monitor the severity of depression and response to treatment.

The app includes notifications to prompt users to log their moods at 9am, 12pm and 8pm. The data used in this study was derived from users who completed scales and mood logs between August 2017 to November 2018.

5

3.2. *Procedures for data analytics*

This study has received ethical approval from the School of Communication and Media Ethics Filter Committee, Ulster University. The data analytics methodology involved using the Health Interaction Log Data Analytics (HILDA) workflow pipeline (*32*), developed from our work in analysis of digital well being behaviour logs (*33*) in various application domains, including telephone crisis lines (*34*, *35*), apps for use by people with intellectual disabilities (*36*) as well as our earlier work in maternal mental health (*37*). The initial stage involved data cleansing.

Standardising workflows is crucial in order to ensure consistency and best practices in a domain. A number of standards for carrying out a data science, data mining or a machine learning project have been proposed. For example, the Cross-industry standard process for data mining (CRISP-DM) is a data mining process model encompassing the following stages: Business understanding, Data understanding, Data preparation, Modeling, Evaluation, Deployment (*38*). CRISP-DM has been available in various guides since 1996. An updated variant developed by IBM, called Analytics Solutions Unified Method for Data Mining/Predictive Analytics (ASUM-DM) (*39*) expands on CRISP-DM.

The HILDA workflow is specifically designed for processing of user, system event and interaction data logged on products and services that use digital well-being interaction technologies. A user event log must at least be comprised of three columns or variables as part of a tabular structure, comprising a unique identifier for the user (can be anonymous); the event that was recorded; and the date and time (preferable precision to seconds) of that event.

The first phase of the HILDA workflow model involves data cleaning. Following data preparation, data prospecting is carried out. Following data prospecting is often a machine learning phase which can involve both unsupervised and supervised machine learning. The HILDA workflow model covers data preparation, data prospecting and machine learning in some detail but not deployment and usage. Once the machine learning stage is complete, the results can lend themselves to being labelled and described in clear terms that are easily understood beyond the research community.

The dataset comprised Comma Separated Values (CSV) and JavaScript Object Notation (JSON). The JSON data, which included mood logs, were transformed into a data frame that was amenable for statistical analysis. The following stage involved the normalisation of dates. The dataset involved different date formats that were reconciled. Dates were also normalised so that each user who started using the app on different dates could be compared or aggregated, e.g. 'day 1' was the relative first day of usage regardless of the absolute date the user started using the app. This was followed by exploratory data analysis (EDA). EDA entailed analysis of probability density functions, histograms, frequency tables and time series analysis to study the temporal aspects of user behaviour. This specifically involved a detailed analysis of the number of mood log completions and the number of scale completions for each hour of the day and for each day of the week. The correlation between these vectors were studied and the outliers based on a regression model were analysed. An investigation into the change of moods over a 24-hour day was also studied as well as the frequency of each mood being logged. Subsequently, a series of features to characterise the behaviour of each user was extracted from the usage dataset to form a new user dataset. These features included, 1) the number of mental health scale completions the user made, 2) the number of mood logs the user made, 3) the user's tenure of using the app (in days), 4) the mean interval between each usage day (in days) and 5) the standard

deviation of the intervals between each usage day. These features were then inputted into k-means clustering, an unsupervised machine learning technique. This technique is used to discover clusters or user groups that are somewhat distinct from one another. The elbow method was used to determine the number of clusters to be computed. Once clusters were determined, we then described these user groups using the cluster centroids which represent the archetypical or average user in that cluster.

All data analysis was carried out using the R programming language and R Studio. Pearson product-moment correlation coefficients were used for association analysis between variables. Significance testing was also computed for correlations, where $p<0.05$ was considered statistically significant.

4. Results

The average app tenure was 5.359 days (SD=26.869 days), however the maximum tenure was 294.00 days. A total of 1,461 users had completed at least one mental health scale and 1,267 users completed at least one mood log. Many new users installed the app and were attracted to the mental health scales but never progressed to use the app or log moods. This is corroborated given that there were 2,129 scale completions and 6,993 mood log completions. The ratio of 3.28 mood log competitions to each scale completion indicates that users prefer EMA in the form of mood logging, but are still willing to complete a mental health scale, albeit infrequently.
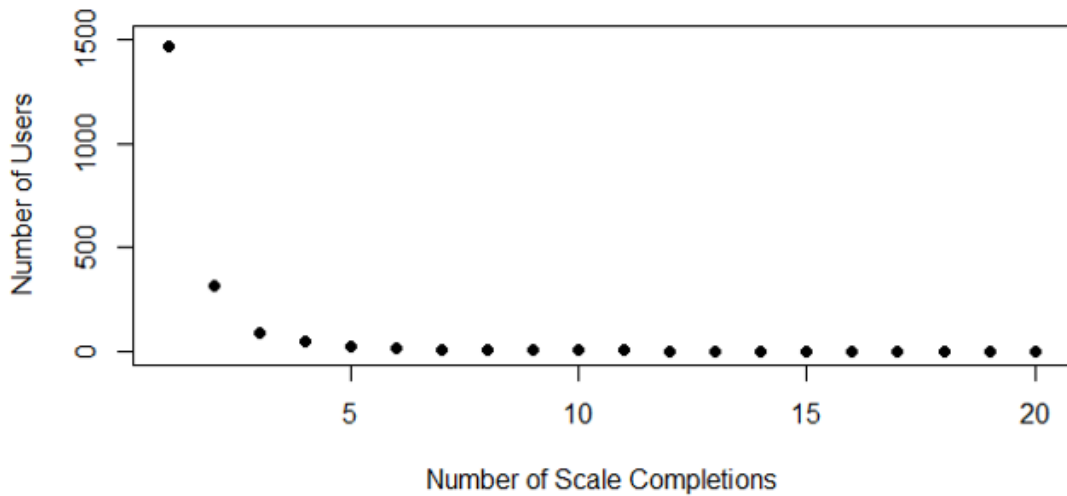

Figure 2. Function between the number of users and the number of mental health scale completions, e.g. 92 users had completed a scale at least 3 times.

Figure 2 shows a function between the number of users and the number of mental health scale completions. There was a total of 1,378 completions of the EPDS scale, 502 completions of the GAD-7 anxiety scale and 248 completions of the PHQ-9 depression scale. However, only 92 users had completed a scale at least 3 times. Figure 2 shows that users who have 4 or 5 scale completions are likely to adopt the app and consistently complete mental health scales. A total of 36 users had completed the EPDS scale at least 3 times. For these 36 users, there is a poor but statistically significant negative correlation (r = -0.23, p-value < 0.01) between the temporal instance number of each EPDS completion and the scale score.
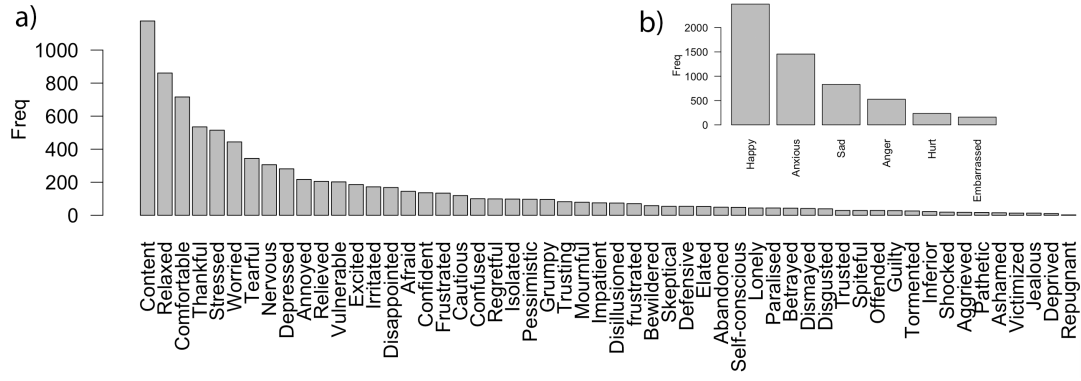
7

Figure 3. a) Frequency of specific mood logs and b) frequency of generic mood logs.

Figure 3 indicates that the most frequent positive generic mood log is 'happy' with the moods 'content', 'relaxed', 'comfortable' and 'thankful' being the most common positive 'specific' moods logged.

Figure 4(a) shows that users are most likely to log moods at 9am, 12pm and 8pm, in line with app prompting. Figure 4 (b) shows that happiness peaks at 1pm. Whilst happiness is the dominant mood during day time, the moods 'anxiety' and 'sadness' peak during night time, with sadness peaking at 1am and anxiety peaking at 4am, and anger peaks at tea-time (5pm and 6pm).

Figure 5 a) and b) shows the frequency of mood logs and scale completions per hour and per day, respectively. The majority of mood logs and scale completions are committed on a Tuesday, Wednesday and Thursday. The fewest number of mood log and scale completions take place on a Saturday followed by a Sunday. Figure 5 also indicates that the hourly patterns of scale completions are different from the hourly patterns of mood log completions. Whilst users do not complete mood logs in the late evening, the hours between 8pm and 10pm are the most common times for users to complete a mental health scale. Users are more willing to log moods at 9am and 12pm as opposed to completing a mental health scale.

For example, Figure 6 shows that users are 8 times more likely to log a mood at 9am than they are to complete a mental health scale. This is why mood log EMA questions are important for collecting highly time-sampled data about well-being, whereas mental health scale EMAs are heavily biased towards evening participation.

Figure 7 shows a) a regression model for the number of mood log completions and scale completions at each hour of the day and b) accompanying residual plot highlighting outliers. There is a strong correlation between the number of mood logs and scale completions per hour (r=0.710, p<0.001). From analysing the residuals, there are more mood logs at 9am and 12pm than would be predicted by the model and significantly more scale completions at 9pm and 10pm than would be predicted by the model, with 10pm having the largest residual from the regression line. This indicates that users are much more willing to complete mental health scales in the late evening, despite the 9am, 12noon and 8pm prompting schedule.

Clustering results are shown in Table 1. The elbow method as shown in Figure 8(a) suggested a four cluster solution that would explain 80.64% of the variability based on the features. A description of these clusters can be seen in Table 2. The clustering model has been visualised in Figure 8(b) using two principal components.

Figure 9 also shows the boxplots for each feature per cluster, and indicates that tenure could the strongest single feature for discriminating between the four clusters.

Table 1. Centroid feature values for each cluster.

Cluster	1	2	3	4
Size	13	4	49	1186
Number of scale completions	2.23	29.5	2.285	0.774
Number of mood log completions	4.615	238.5	35.469	3.433
Tenure (days)	176.615	133.25	98.877	2.289
Mean interval between use	42.071	0.812	8.377	0.407
SD of intervals between use	76.274	2.243	15.768	0.606

Table 2. Cluster descriptions.

| Cluster	Description
1	This small group (1% of users) use the app for almost 6 months but only complete 2 scales and log 4 or 5 moods. They can go for over a month without using the app and their next interaction is unpredictable
2	This very small group (0.319% of users) are high adopters and frequent users. They use the app for over 4 months. Each user is willing to complete almost 30 scales and log over 200 moods. There are very frequent users that log their mood multiple times in a given day.
3	This group (3.913% of users) use the app for over 3 months completing 2 scales and logging 35 moods. The interval between each mood log is on average 8 days, however this interval can greatly vary
4	Like many apps, the majority of these users (94.728%) are transient users and not adopters (perhaps due to app fatigue). This group only uses the app for 2 or 3 days and only log 3 moods. They do log their moods multiple times during these days but show very little activity, some of which do not complete a scale.
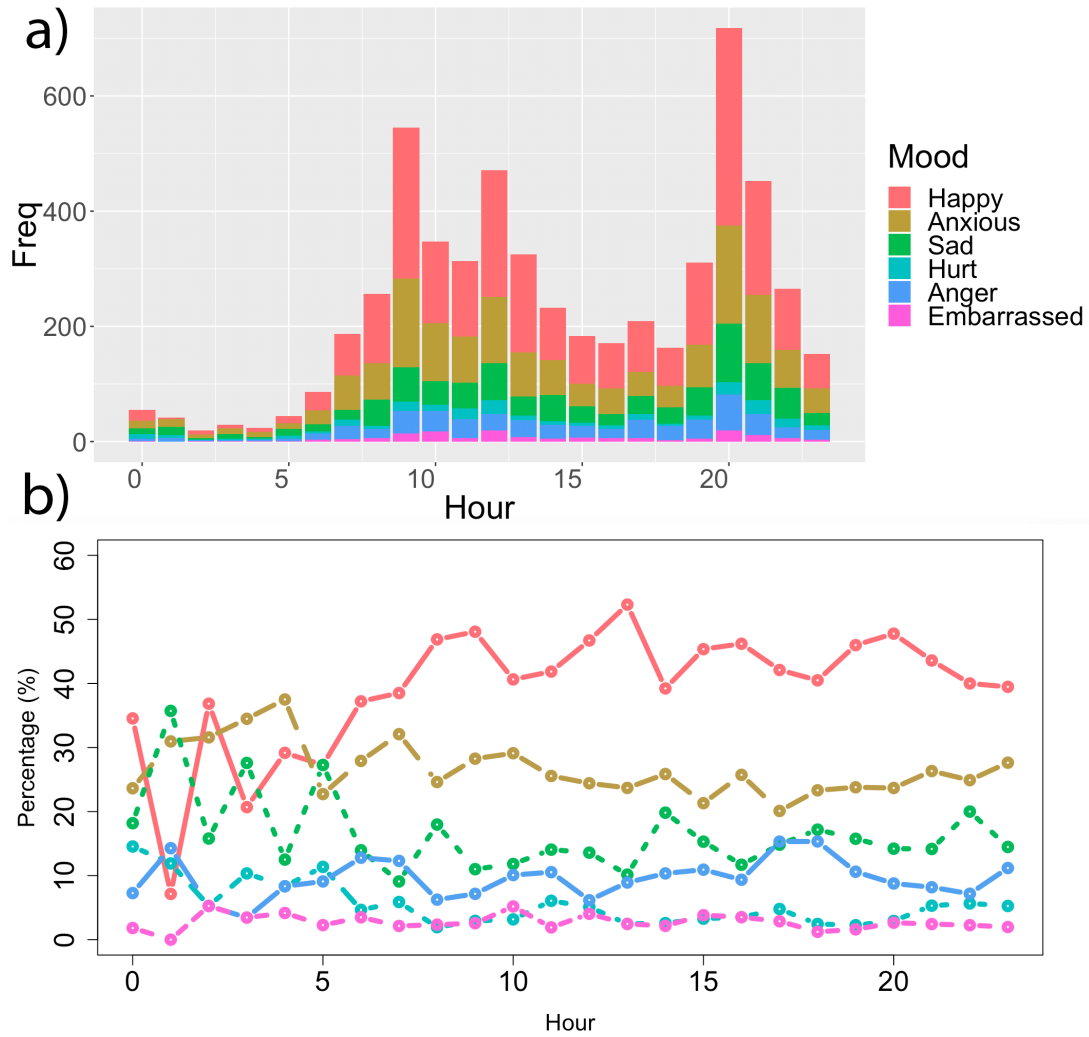
Figure 4. a) Frequency of mood logs per hour of the day and b) Percentage of particular moods being logged at each hour.

5. Discussion

This study used data analytics to investigate the temporal behaviour of users when completing mental health scales and mood logs in the form of EMAs using a maternal mental health app. It found that the overall total number of completions was higher for mood logs than scales. Users tend to complete logs at the time of the notifications but are much more willing to complete scale surveys in the late evening (9-10pm).

The most frequently reported mood was 'happy', which was the dominant mood during day time, however 'anger' peaks at tea-time (5pm and 6pm) and the moods 'anxiety' and 'sadness' peak during night time, at 1am and 4am respectively. The majority of completions were on Tuesday, Wednesday and Thursday, with the fewest number of logs being at the weekend, especially on Saturdays. There were superficially more unique users who completed a scale in contrast to users who completed a mood log, as many new users installed the app and were attracted to the mental health scales but never progressed to sustained use of the app or mood logging.
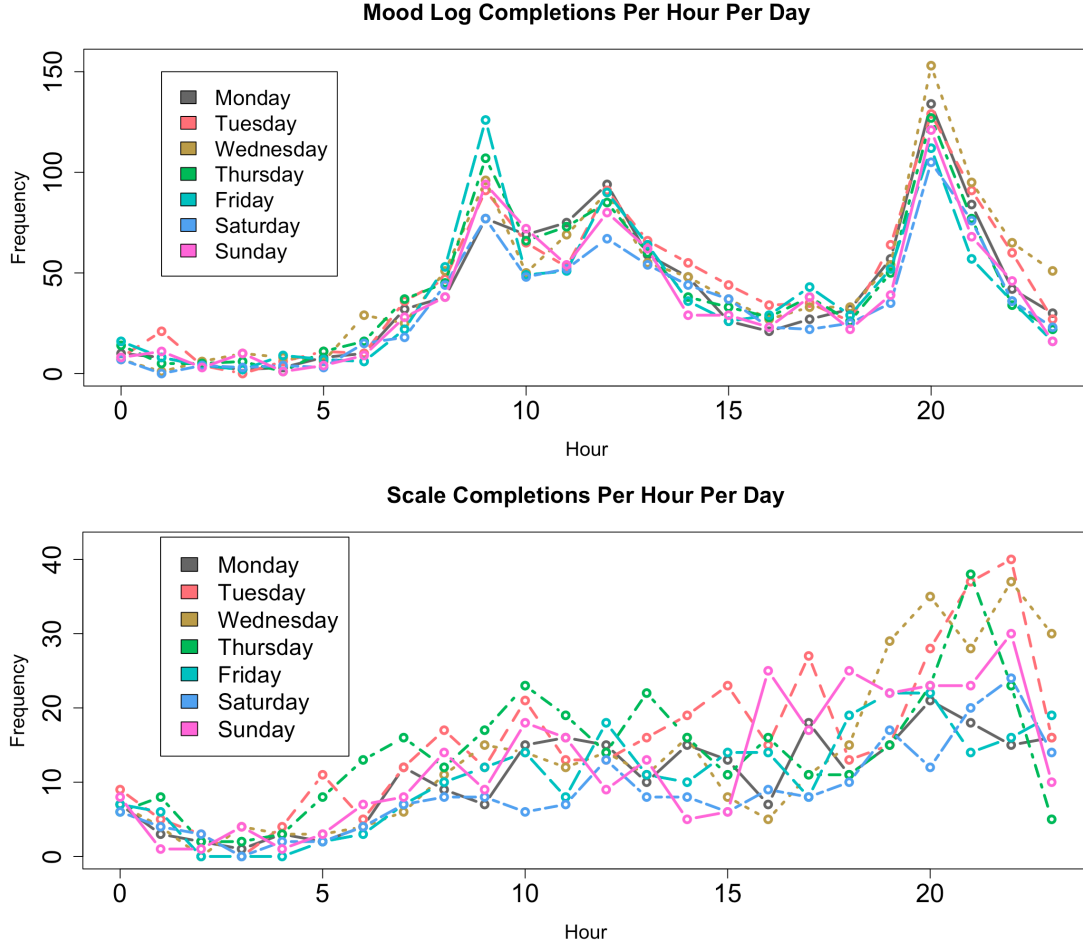
Figure 5. a) Frequency of mood logs and b) Mental health scale completions per hour & day.

From the 4 cluster groups indicated from the data, the majority of these users (94.728%) are light users and not adopters, which is similar to other health apps and akin to typical app retention graphs.

For 36 users, there was a poor but statistically significant negative correlation between the temporal instance number of each EPDS completion and the mental health scale score. This suggests that postnatal depression decreases for some users as they continue using the app. It is possible that users are more likely utilise local resources and seek external help as a result of logging moods and using the locator and support tools within the app, however further work is needed to confirm this.

These results indicate that it is important to encourage users to complete scales in order to secure them as full adopters of the app. When an app user completes four or five mood log or scale completions, they adopt the app and continue to use it. If engagement of the app can be increased over a short period of time, it is more likely to create a behavioural change and result in the app being part of the users routine (*40*). User notifications have an important role in encouraging use of the app. Given that late evening (9-10pm) is the most common time to complete a log or scale, an additional notification could be released at 9pm to encourage completion of EMAs. Personalisation is another method for increasing engagement with apps, such as including the users name in push notifications. Tailoring an app for individual needs
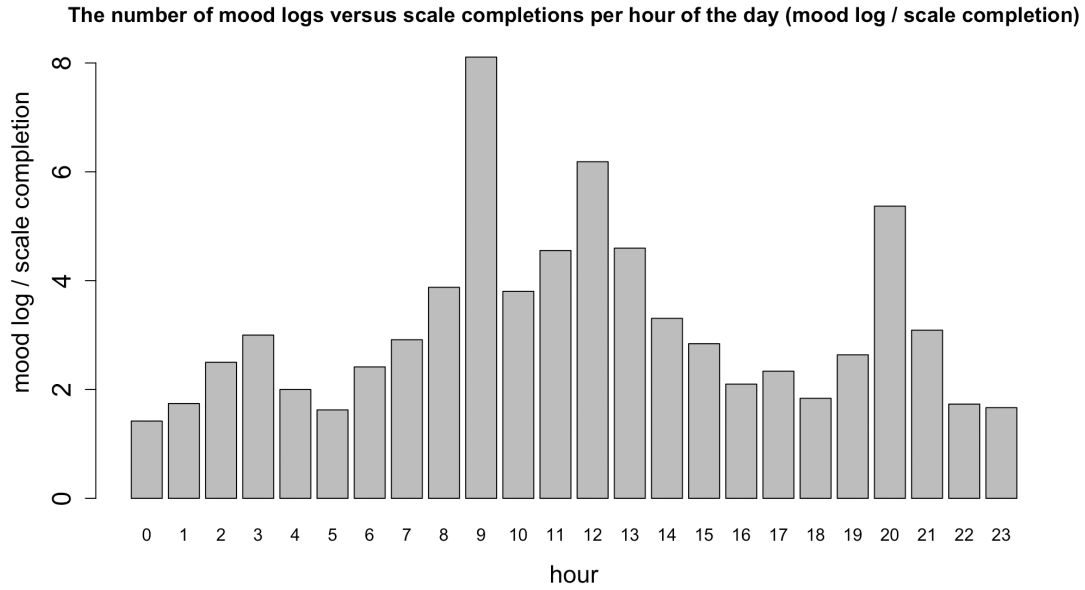
The number of mood logs versus scale completions per hour of the day (mood log / scale completion)

Figure 6. Number of mood log completions versus mental health scale completions per hour of the day (number of mood log completions / number of scale completions).

is important for user retention, which can include displaying information and prompts based on work schedules, sleep and location triggers (*41*).

The questionnaires and scales that are used within the app are routinely administered in more formal clinical settings. Typically, this makes it difficult to reach those who are not actively engaged in health care (*42*). To resolve this, a smartphone app in which users complete EMAs can be used to reach more vunerable groups such as new mothers to support self-monitoring of mental health.

It is clear that using digital technologies can enhance healthcare by ensuring more people have access to high-quality mental health treatment, advice and support. In the fields of psychiatry and clinical psychology, digital phenotyping has huge potential to improve the quality of care and services worldwide.

6. Conclusion

This study used a subset of digital phenotyping known as EMA to analyse behaviour and user engagement of two EMA approaches - one model that uses mental health scales, which are time consuming and cognitively demanding for users and the other EMA model comprised of mood logs, which are more efficient to complete requiring less mental workload. We can conclude that users prefer mood log EMAs and that the temporal behavior of users engaging with EMAs in the form of mental health scales are distinctly different from how they engage with mood logs.
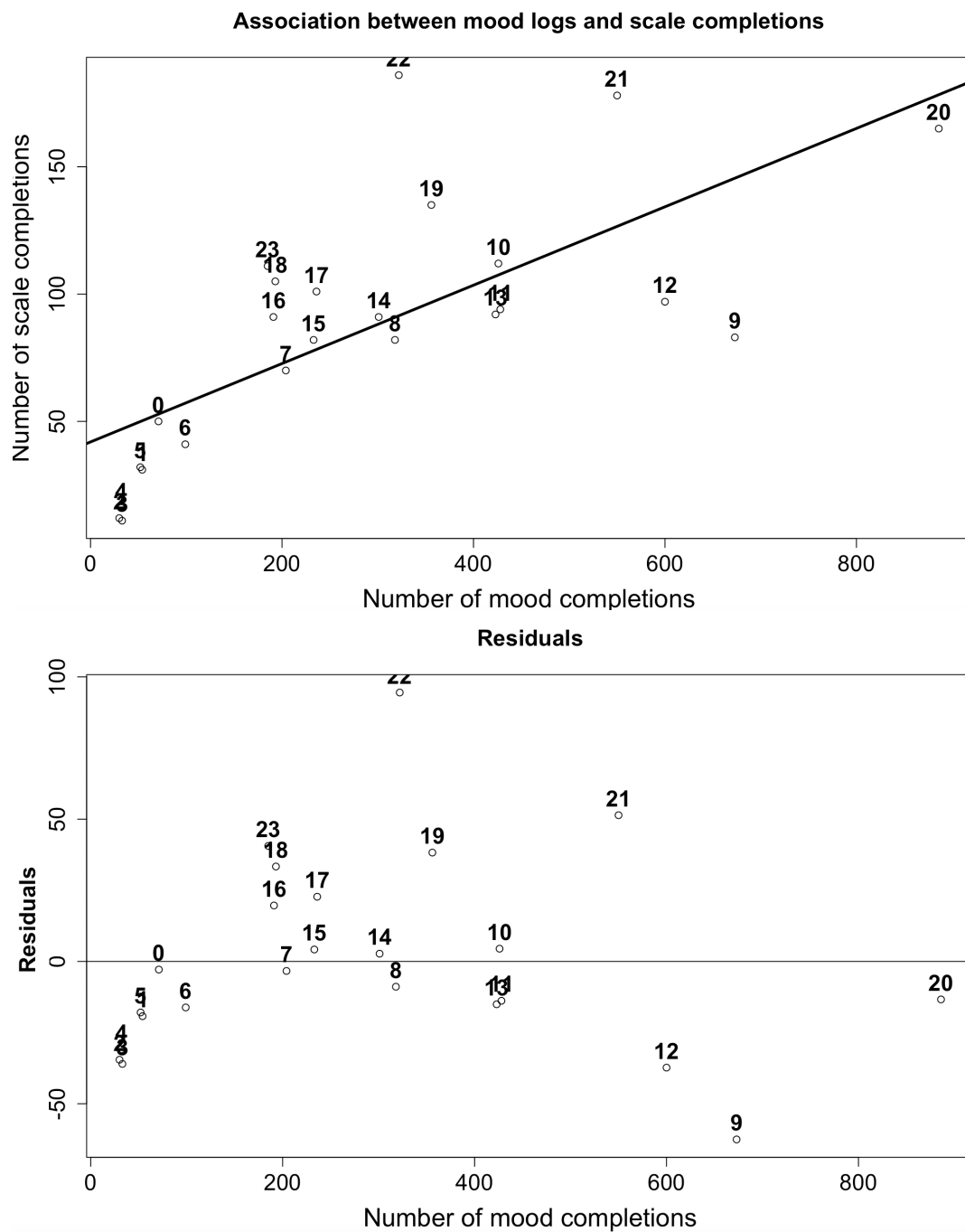
Acknowledgement(s)

Figure 7. a) The association between the number of mood log completions and scale completions at each hour of the day (hour is labelled beside each point and b) residual plot.
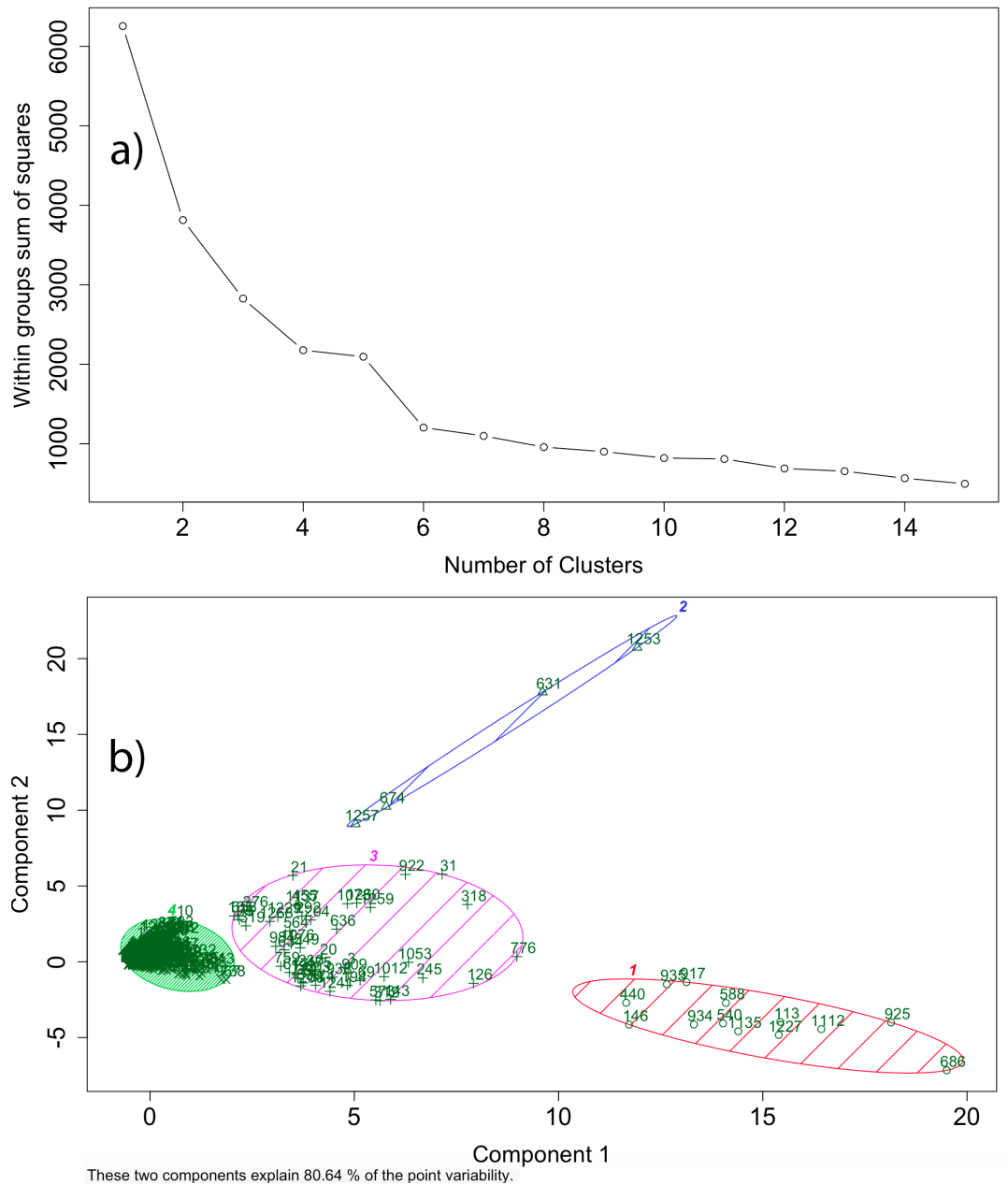
Figure 8. a) the elbow plot showing within group sum of squares per cluster solution, b) the selection cluster solution (n=4) as depicted using two principal components.
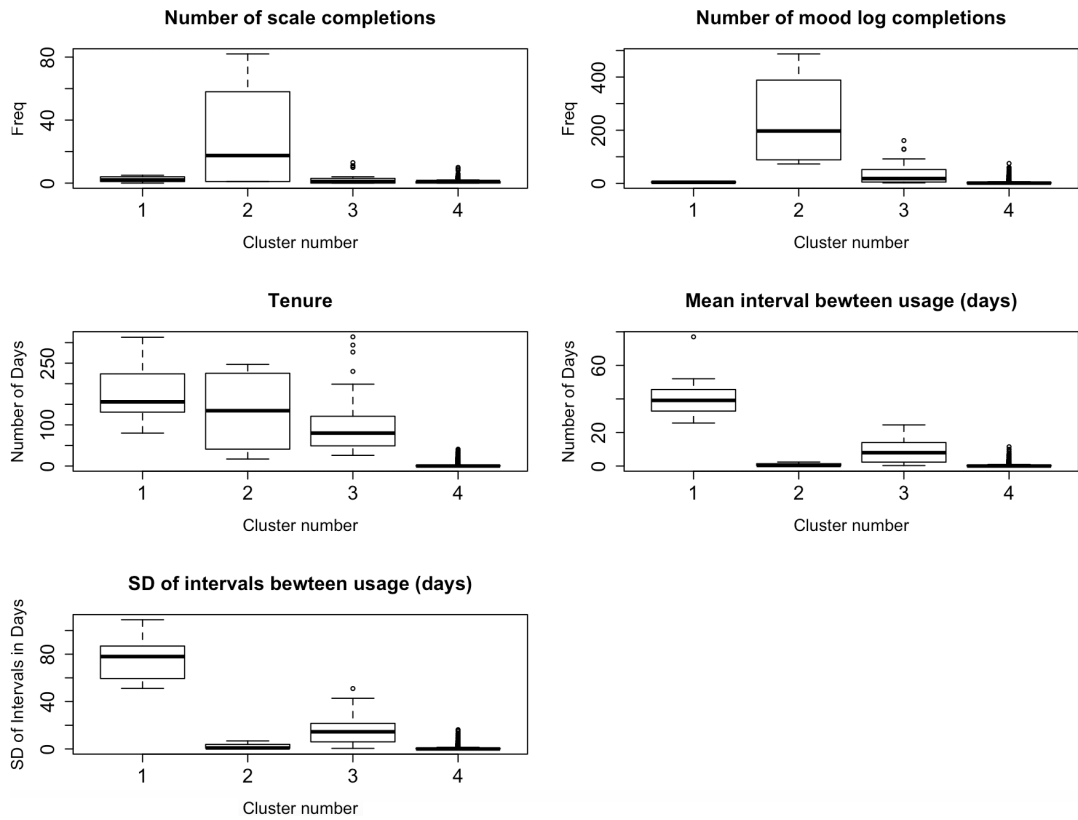
15

Figure 9. Boxplots for each feature per cluster.

References

(1) World Health Organisation. Maternal mental health **2019**. Accessed 18 February 2019, `http://www.who.int/maternal`.

(2) Maternal Mental Health Alliance. Maternal mental health **2017**. Accessed 11 May 2018, `https://maternalmentalhealthalliance.org`.

(3) National Health Service. Maternal mental health **2017**. Accessed 18 February 2019, `https://www.nhs.uk/conditions/pregnancy-and-baby/mental-health-problems-pregnant/`.

(4) Bauer, A.; Parsonage, M.; Knapp, M.; Iemmi, V.; Adelaja, B. Costs of Perinatal Mental Health Problems **2014**. Accessed 18 June 2019, `https://www.nwcscnsenate.nhs.uk/files/3914/7030/1256/Costs_of_perinatal_mh.pdf`.

(5) Royal College of Midwives. Maternal mental health Improving emotional wellbeing in postnatal care. Pressure Points **2014**.

(6) Parfitt, Y.; Ayers, S. Postnatal mental health and parenting: The importance of parental anger., *Infant mental health journal* **2012**, *33* (4), 400–410.

(7) Etherington, N.; McDougall, J.; DeWit, D.; Wright, V. Maternal factors and the emotional and behavioural functioning of adolescents with chronic health conditions, *Disability and Rehabilitation* **2016**, *38* (14), 1359–1369.

(8) Royal College of Obstetricians and Gynaecologists. Maternal mental health - womens voices **2017**. Accessed 1 March 2019, `https://www.rcog.org.uk/globalassets/documents/patients/information/maternalmental-healthwomens-voices.pdf`.

(9) National Institute for Health and Care Excellence. Antenatal and postnatal mental health **2016**. Accessed 1 March 2019, `https://www.nice.org.uk/guidance/qs115/resources/antenatal-and-postnatal-mental-health-pdf-75545299789765`.

(10) Miller, K.; Woollam, P.J.; Powell, G.; Hitchings, D.; Stallard, J. A rehabilitation device data logging system, *Disability and Rehabilitation: Assistive Technology* **2007**, *2* (1), 9–14.

(11) de Santana, V.F.; Baranauskas, M.C.C. Summarizing Observational Client-side Data to Reveal Web Usage Patterns, In *Proceedings of the 2010 ACM Symposium on Applied Computing*; ACM, 2010; pp 1219–1223. `http://doi.acm.org/10.1145/1774088.1774344`.

(12) Woo, D.; Mori, J. *Accessibility: A Tool for Usability Evaluation*, Masoodian, M., Jones, S., Rogers, B., Eds.; Springer-Verlag Berlin Heidelberg, 2004; pp 531–39.

(13) Morrison, C.; Doherty, G. Analyzing engagement in a web-based intervention platform through visualizing log-data., *Journal of medical Internet research* **2014**, *16* (11), e252.

(14) Torous, J.; Kiang, M.V.; Lorme, J.; Onnela, J.P. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research, *JMIR Mental Health* **2016**, *3* (2), e16.

(15) Onnela, J.P.; Rauch, S.L. Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health, *Neuropsychopharmacology* **2016**, *41* (7), 1691–1696.

(16) Martinez-Martin, N.; Insel, T.; R Dagum, P.; Greely, H.T.; Cho, M.K. Data mining for health: staking out the ethical territory of digital phenotyping, *npj Digital Medicine* **2018**, *1* (1), 68.

(17) Bell, I.H.; Fielding-Smith, S.F.; Hayward, M.; Rossell, S.L.; Lim, M.H.; Farhall, J.; Thomas, N. Smartphone-based ecological momentary assessment and intervention in a blended coping-focused therapy for distressing voices: Development and case illustration., *Internet Interventions* **2018**, *14*, 18–25.

(18) Insel, T.R. Digital phenotyping: a global tool for psychiatry, *World Psychiatry* **2018**, *17* (3), 276–277.

(19) Lewin, K. *A dynamic theory of personality: selected papers*; McGraw-Hill, 1935.

(20) Larson, R.; Csikszentmihalyi, M. The Experience Sampling Method. In *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi*; Dordrecht: Springer Netherlands, 2014.

(21) Shiffman, S.; Stone, A.; Hufford, M. Ecological Momentary Assessment, *Annual Review of Clinical Psychology* **2008**, *4* (1), 1–32.

(22) Place, S.; Blanch-Hartigan, D.; Rubin, C.; Gorrostieta, C.; Mead, C.; Kane, J.; Marx, B.; Feast, J.; Deckersbach, T.; Pentland, A.; et al. Behavioral Indicators on a Mobile Sensing Platform Predict Clinically Validated Psychiatric Symptoms of Mood and Anxiety Disorders., *Journal of Medical Internet Research* **2007**, *19* (3), e75.

(23) Intille, S.S.; Rondoni, J.; Kukla, C.; Ancona, I.; Bao, L. A Context-aware Experience Sampling Tool, In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA; ACM, 2003; pp 972–973. `http://doi.acm.org/10.1145/765891.766101`.

(24) Mulder, I.; Hofte, G.T.; Kort, J. SocioXensor: Measuring user behaviour and user eXperience in conteXt with mobile devices **2005**.

(25) Wiebe, D.; Nance, M.; Houseknecht, E.; Grady, M.; Otto, N.; Sandsmark, D.; Master, C. Ecologic Momentary Assessment to Accomplish Real-Time Capture of Symptom Progression and the Physical and Cognitive Activities of Patients Daily Following Concussion, *JAMA Pediatrics* **2016**, *170* (11), 1108–1110.

(26) Colombo, D.; Palacios, A.G.; Alvarez, J.F.; Patane, A.; Semonella, M.; Cipresso, P.; Kwiatkowska, M.; Riva, G.; Botella, C. Current state and future directions of technology-based ecological momentary assessments and interventions for major depressive disorder: protocol for a systematic review, *BMC Systematic Reviews* **2018**, *7* (233), 1108–1110.

(27) Álvarez, R.; Murua, A.; Artetxe, A.; Epelde, G.; Beristain, A. A Platform for User Empowerment Through Self Ecological Momentary Assessment / Intervention, In *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare*; ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2015; pp 206–209.

(28) Artetxe., A.; Epelde., G.; Beristain., A.; Murua., A.; lvarez., R. Gaining Insight from Physical Activity Data using a Similarity-based Interactive Visualization, In *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 2: IVAPP, (VISIGRAPP 2016)*; SciTePress, 2016; pp 115–122.

(29) Cox, J.; Holden, J.; Sagovsky, R. Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale, *British Journal of Psychiatry* **1987**, *150*, 782–786.

(30) Goldberg, D. *Manual of the General Health Questionnaire*; NFER-Nelson, 1978.

(31) Kroenke, K.; Spitzer; Williams, J. The PHQ-9: validity of a brief depression severity measure, *Journal of General Internal Medicine* **2001**, *16* (9), 606–13.

(32) Mulvenna, M.; Bond, R.; Grigorash, A.; O'Neill, S.; Ryan, A. HILDA - A Health Interaction Log Data Analysis Workflow to Aid Understanding of Usage Patterns and Behaviours. In *Proceedings AISB*; AISB, 2018.

(33) Turkington, R.; Mulvenna, M.; Bond, R.; O'Neill, S.; Armour, C. The Application of User Event Log Data for Mental Health and Wellbeing Analysis, *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI-2018)* **2018**.

(34) ONeill, S.; Bond, R.R.; Grigorash, A.; Ramsey, C.; Armour, C.; Mulvenna, M.D. Data analytics of call log data to identify caller behaviour patterns from a mental health and well-being helpline, *Health Informatics Journal* **2018**, 1460458218792668.

(35) Grigorash, A.; O'Neill, S.; Bond, R.; Ramsey, C.; Armour, C.; Mulvenna, M. Predicting Caller Type From a Mental Health and Well-Being Helpline: Analysis of Call Log Data **2018**, *5* (2).

(36) Donnelly, M.; Bond, R.; Mulvenna, M.; Taggart, L.; Hill, D.; Fitzsimons, P.; Martin, S.; Hassiotis, A. *Facilitating Social Connectedness for People with Autism and Intellectual Disability Using an Interactive App*, Bond, R., Mulvenna, M., Wallace, J., Black, M., Eds., 2018.

(37) Moorhead, A.; Bond, R.; Mulvenna, M.; O'Neill, S.; Murphy, N. *A Self-Management App for Maternal Mental Health*, Bond, R., Mulvenna, M., Wallace, J., Black, M., Eds., 2018.

(38) Shearer, C. The CRISP-DM Model: The New Blueprint for Data Mining, *Journal of Data Warehousing* **2000**, *5*.

(39) Haffar, J. Have you seen ASUM-DM? **2015**. Accessed 1 March 2019, `ftp://ftp.software.ibm.com/software/data/sw-library/services/ASUM.pdf`.

(40) Torous, J.; Nicholas, J.; Larsen, M.E.; Firth, J.; Christensen, H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements, *Evidence-Based Mental Health* **2018**, *21* (3), 116–119.

(41) Srinivas, P.; Bodke, K.; Ofner, S.; Keith, N.; Tu, W.; Clark, D. Context-Sensitive Ecological Momentary Assessment: Application of User-Centered Design for Improving User Satisfaction and Engagement During Self-Report, *JMIR Mhealth Uhealth* **2019**, *7*, e10894.

(42) Haberer, J.; Trabin, T.; Klinkman, M. Furthering the reliable and valid measurement of mental health screening, diagnoses, treatment and outcomes through health information technology., *General Hospital Psychiatry* **2013**, *35*.