Lee, James Alexander (2018) ***Exploring the Use of Online Social Network Activity and Smartphone Photography as an Intervention to Track and Influence Emotional Well-Being.*** **Doctor of Philosophy (PhD) thesis, University of Kent,.**

Exploring the Use of Online Social Network Activity and Smartphone Photography as an Intervention to Track and Influence Emotional Well-Being

A thesis submitted to the University of Kent

for the degree of Doctor of Philosophy in Digital Arts

By

James Alexander Lee

September 2018

> “Two groups of stakeholders that collaborate to develop technology supported behavioral change interventions are behavioral scientists and computer science engineers.
>
> ...
>
> A fruitful and needed outcome of such interdisciplinary cross-talk should be to bring 20th-century behavioral theories up to the task of guiding intervention delivery in the age of mobile technologies and mobile-enabled interventions.”

— Spring et al. [173]

Abstract

The proliferation of internet and mobile technologies has expanded the means of detecting and influencing mental health, with this thesis focusing on the affective phenomena associated with emotional well-being including mood, affect and emotion. Traditional detection techniques including surveys and self-reports are grounded in the psychological literature; however, they introduce an inhibiting burden on the participants. The ability to passively detect psychological state using technologies including online behavioural tracking and mobile sensors is a prevalent focus of the current literature. Traditional positive psychology interventions commonly involve emotionally expressive writing tasks which can also be tedious for participants. Augmenting traditional intervention techniques with technologies such as smartphone applications can be one method to modernise interventions.

The first research study in this thesis aimed to utilise online social network (OSN) activity to detect mood changes. The study involved collecting the participants' behavioural activities such as likes, comments and tweets from their Facebook and Twitter profiles. Machine learning was used to create an algorithm to classify participants according to their online activity and their self-reported mood as ground truth. The findings indicated that participants can be grouped into those who displayed positive, negative or weak correlations with their online activity. Following the classification, the system used a sliding window of 7 days to track the participant's mood changes for those in the positive and negative groups.

The second research study introduced a positive psychology intervention in the form

of a smartphone application called SnapAppy which promotes positive thinking by integrating momentary smartphone photography with traditional intervention methodologies. Participants were required to take photos and write about positive moments, past events, acts of kindness and gratuitous situations, encouraging them to think more positively. The results indicated that features such as the number of photos taken, the effort applied to annotating the photos, the number of photos revisited and the photos containing people were positively correlated with an improvement in mood and affect.

The product of this thesis is a novel method of passively tracking mood changes using online social network activity and an innovative smartphone intervention utilising photography to influence emotional well-being.

Acknowledgements

Firstly, I would like to express my sincere thanks to my supervisor, Dr Christos Efstratiou, for his exemplary guidance, advice, support and encouragement throughout my PhD journey. His ability to see the light at the end of the tunnel, even during the toughest obstacles, was invaluable. I would also like to thank my second supervisor, Dr Jim Ang, for his additional support. Thanks for being so enthusiastic about everything.

I am grateful to Lu Bai and Panote Siriaraya who provided their expertise for parts of my analysis, without them I would still be staring blankly at my spreadsheet of data. I am also very grateful to Professor Roger Giner-Sorolla, Dr Dinkar Sharma, Dr Ana Carla Crispim and Ben Tappin from the School of Psychology at the University of Kent for their assistance with all psychology-related matters. Thanks also to the School of Psychology's RPS team for providing access to their recruitment system.

A special thanks goes to James Rowe for dedicating his time to design the SnapAppy logo. Thanks are also due to Boris Otkhmezuri, Chiara Lunerti, Jake Minns, James Rowe, Jon Baker, Jittrapol Intarasirisawat, Pruet Putjorn, Rina Mashimo and Suzy Parker for their fantastic feedback while I was developing SnapAppy and during the user trial. I would also like to express my gratitude to all the participants who took part in my studies. I hope that those who participated in the SnapAppy study derived some joy from the app, however small.

I would like to thank my friends and colleagues from the School of Engineering and Digital Arts at the University of Kent for providing a friendly work environment and for keeping me sane during these four years: Jon Baker, Deogratias Mzurikwao, Rina Mashimo, Max Tayler, Jittrapol Intarasirisawat, Hamed Alsufyani, Michael Green, Luma Tabbaa, Boris Otkhmezuri, Chiara Lunerti, Sumita Chauhan, Sara Choudhrey, Pruet Putjorn, Ben Nicholls and Zoe Mairs.

Finally, for their continued support and belief in me, my profound thanks goes to Chris Carter, Lee Ward, Darren Horner, Reece Mcilwraith, Matthew Banks, Kelly Tung, Rebecca Tizzard, Jacob Hammond, Joseph Williams, Theologis Strikos and to my family: Mum, Dad and all those we have lost.

Contents

List of Figures

List of Tables

List of Acronyms

AHI	Authentic Happiness Inventory
API	Application Programming Interface
CBT	Cognitive Behavioural Therapy
DES	Differential Emotion Scale
DRM	Day Reconstruction Method
ECG	Electrocardiogram
EEG	Electroencephalogram
EMA	Ecological Momentary Assessment
EMG	Electromyogram
FSS	Fear Survey Schedule
GALC	Geneva Affect Label Coder
GPS	Global Positioning System
GSR	Galvanic Skin Response
LIWC	Linguistic Inquiry and Word Count
OSN	Online Social Network
PAM	Photographic Affect Meter
PANAS	Positive and Negative Affect Schedule
POMS	Profile of Mood States
PWB	Psychological Well-Being
SAM	Self-Assessment Manikin
SHS	Subjective Happiness Scale
STAS	State-Trait Anger Scale
SWB	Subjective Well-Being
SWLS	Satisfaction with Life Scale
TOSCA	Test of Self-Conscious Affect
UI	User Interface

— 1 —

Introduction

1.1 Detecting and Influencing Emotional Well-Being

Mental health problems are one of the main causes of disease burden worldwide [189]. Considering the impact that mental health can have in society and the relatively high cost of offering appropriate support, significant effort is being made by the research community to explore new techniques for diagnosis, progress tracking and intervention to improve various aspects of mental health. As defined in Section 2.1, emotional well-being is an umbrella term encompassing a variety of interconnected aspects relating to one's mental health including well-being, mood, emotion and affective state. Mood, emotion and affect are affective phenomena which can contribute to mental health problems and will be the focus of this thesis.

The traditional methods of psychological assessment commonly involve surveys, assessments and diaries in order to log an individual's affective state during their day [15, 70, 167]. These methods have been developed to accurately assess one's affective state in the majority of cases; however, they are not devoid of limitations. These types of assessment can induce burden on the participants and can disrupt their daily activities. The data can also be compromised by falsification, non-compliance, satisficing and confusion due to linguistic and cultural issues. Thus, research into the passive detection of affective states may be one solution to these limitations.

Passive detection technologies can include using sensors in smartphones and wearables to infer the psychological state of the user through their physical location [158], social activities [81] and behavioural habits [87]. Online social networks have also been the focus of research into passive detection of emotional well-being, using behavioural patterns and sentiment analysis to predict the affective state of their users and is studied further in Chapter 3. The major limitation of the research in this field is that the inferred affective states are rarely verified against real-world ground truth data. Researchers rely on the online behaviours of individuals from a single social network to be truthful, exhaustive and an accurate representation of their real-world feelings. This limitation motivated the work in Chapter 3.

Logging, assessing or inferring psychological state provides the data about one's emotional well-being; however, without being interpreted and acted upon it does not have any impact on one's life. The study in Chapter 3 provides only a case study method for passively extracting emotional well-being information; the data is not visualised nor acted upon. People can be supported to consciously think about their emotional well-being in order to make a positive change. Positive psychology interventions are defined as "treatment methods or intentional activities aimed at cultivating positive feelings, positive behaviors, or positive cognitions" [169] with the goal of making improvements to one's happiness, subjective well-being, satisfaction with life, flourishing, mindfulness or positive thinking [163]. Interventions are one of many methods to make a positive change to one's life by influencing aspects of one's psychological state. Influential methodologies traditionally involved writing or thought-provoking tasks focused on positive and emotional experiences [18, 128], reminiscence [24, 27], kindness [120], gratitude [43, 97] and goal setting [72, 102, 165].

The widespread usage of modern smartphones has enabled traditional intervention methodologies to be applied digitally and remotely. The portability, passive sensing, notifications and online features are a number of advantages to modern interventions. The use of smartphone and internet technologies for positive psychology interventions has thus become more pervasive in this field of research and the focus is on using these technologies to explore new opportunities for technology-enhanced interventions. One of the common activities that smartphone users carry out on a daily basis is momentary photography, especially with the popularity of apps such as Instagram and Snapchat. Leveraging this common practice of taking and reviewing photos as a positive psychology intervention is the focus of the study in Chapter 4.

Understanding one's emotional well-being, either by manually tracking or passively detecting their psychological state, is a core aspect of providing support to that person. Support might come in various forms including counselling, lifelogging or an intervention. The ability to track and detect psychological state allows the support to provide the appropriate assistance depending on the specific situation. For example, it is important to identify vulnerable individuals who may be suffering from a mental illness or those who are experiencing a depressive episode or a dip in mood in order to trigger a pertinent intervention at an opportune time. For those interested in lifelogging, or for patients who are required to log their activities, it is important that they have the ability and technology to easily track and analyse their behavioural patterns in order to make informed improvements to their life. The development of self-awareness of one's emotional well-being, whether positive or negative, is an important catalyst for such life changes and can be facilitated by passive mood detection. Thus the link between psychological assessments, active tracking, passive detection and interventions form an important relationship where the methodologies and technologies must work together to provide a pragmatic system to improve emotional well-being. This thesis explores two novel case studies focusing on using technologies and methodologies which have not yet been explored in sufficient depth in the fields of passive detection and positive psychology intervention.

1.2 Research Questions

Following the review of literature regarding the passive detection of emotional well-being, it was evident that researchers often relied upon behavioural data from online social networks as being a true representation of one's real-world psychological state [16, 59, 107, 119, 139]; few would validate their findings against real-world ground truth data. It was also found that many papers in this field relied on only a single social network for their data. Some researchers have cautioned over the reliance on a single data source as social network users tend to gravitate towards specific social networks for specific purposes [153]. Thus one of the aims of this thesis is to investigate a novel method of passively detecting affective phenomena using online behavioural data from multiple social networks where the result is validated against real-world ground truth. The following research questions are addressed in Chapter 3:

- **RQ1:** Which specific behavioural activities on Facebook and Twitter correlate with real-world mood changes?
- **RQ2:** Can activity from both Facebook and Twitter be used to infer real-world mood changes?

Existing literature demonstrates extensive use of smartphones for positive psychology interventions to influence emotional well-being. However, there was a large gap in the integration of photography and positive psychology. There is very little research into the positive consequences which result from smartphone photography and the reviewing of photos related to positive experiences. Thus the following research questions are addressed in Chapter 4:

- **RQ3:** Do smartphone photography activities correlate with changes in mood, affect and satisfaction with life?

- **RQ4:** Which specific smartphone photography activities show a statistically significant correlation with change in mood, affect and satisfaction with life?

1.3 Thesis Structure

This thesis contains the following chapters:

Chapter 2 presents a literature review exploring the various aspects of emotional well-being including affective state, emotion, mood, wellness and mental health conditions followed by the traditional and digital assessment methodologies for capturing these affective phenomena. The challenges of these manual assessment methodologies then motivate the review of the technologies used to passively detect different aspects of emotional well-being. Specifically, smartphone sensing, wearable sensing and online behavioural tracking methods are discussed. Next, the research in the positive psychology intervention field is discussed, focusing again on the traditional and digital forms of intervention in the literature. Finally, the real-world applications of these passive detection and influential technologies are summarised, including lifelogging, visualisation and affective computing.

Chapter 3 presents a study exploring the use of online social network activity as an indicator of mood changes. This research combines online activity from both Facebook and Twitter in order to passively infer an individual's real-world mood changes. This work employs experience sampling to capture the mood of the participants as ground truth, rather than relying on an inferred mood state from the online social network data. The results identify behavioural features which are key to the relationship between online activity and real-world psychological state. These features are used to build a classifier which can predict mood changes within a window of 7 days.

Chapter 4 presents a positive psychology intervention involving momentary smartphone photography to influence emotional well-being. The study saw 74 participants use a smartphone application called SnapAppy for one month, taking and reviewing photos related to positive moments, events and experiences. The analysis identifies key behavioural features within the intervention which show correlations with mood and affect. The implications of this research for future interventions, emotional integration with existing photo platforms, passive detection and recommendation systems are discussed.

Chapter 5 presents a summary of the work in this thesis and conclusions are made on the research questions and contributions. Finally, the limitations of the studies and directions for future research are discussed.

1.4 List of Publications

The research in this thesis has been presented, published, or is in the process of being published, in peer-reviewed journals and conferences. These publications are as follows:

Chapter 3:

- J. A. Lee, C. Efstratiou, and L. Bai, "OSN Mood Tracking: Exploring the Use of Online Social Network Activity as an Indicator of Mood Changes," in International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. ACM, 2016, pp. 1171–1179.
- J. A. Lee and C. Efstratiou, "Exploring the Use of Online and Mobile Data to Passively Track and Analyse Psychological State," poster presented at the EDA School Research Conference, 2016.

Chapter 4:

- J. A. Lee, C. Efstratiou, P. Siriaraya, D. Sharma and C. S. Ang, “SnapAppy: A Positive Psychology Intervention using Smartphone Photography to Improve Mental Well-Being,” in Pervasive and Mobile Computing. Elsevier, 2019 [under review].
- J. A. Lee and C. Efstratiou, “SnapAppy: A Positive Psychology Intervention using Smartphone Photography to Improve Mental Well-Being,” poster presented at the EDA School Research Conference, 2018.

— 2 —

Literature Review

This chapter reviews the existing literature regarding the understanding of emotional well-being, the terms used to describe different affective states and the methodologies and surveys used for psychological assessments. This is followed by a review of the existing research focused on tracking and monitoring emotional well-being along with positive psychology interventions used to influence one's emotional state. Lastly, there will be a discussion of the real-world applications of tracking, monitoring and influencing emotional well-being.

2.1 Emotional Well-Being

Mental, psychological and emotional well-being are umbrella terms encompassing a variety of interconnected aspects relating to one's mental health, well-being, mood, emotion and affective state. Terms like "mood" and "emotion" are often used interchangeably, especially by non-academics, which causes their definitions to become unclear. Researchers and psychologists have therefore sought to define clear classifications of these affective phenomena [40, 126]. To begin exploring this broad area of psychology, the current definitions and distinctions of these terms will be established.

2.1.1 Affect

Affect, sometimes known as *core affect*, is the "neurophysiological state consciously accessible as the simplest nonreflective feelings evident in moods and emotions" [150]. It is also described as the way one's psychological state is portrayed, known as *affect* *display*. A person is constantly experiencing affect but the intensity can vary over time. At any given moment, affect is most commonly defined with a positive and negative scale or with a combination of two dimensions: valence and arousal. The valence scale, also known as *pleasure*, can range from distress to ecstasy. The arousal scale, also known as *activation* or *energy*, can range from sleepy, drowsiness or low energy to alertness, excitement or high energy. These dimensions are best visualised on Russell's affect grid [152] (Figure 2.1) and circumplex model of affect [151] (Figure 2.2). Dominance is a third dimension sometimes associated with affective state which represents the amount of control the participant has in the situation being measured and is discussed further in Section 2.2.2.

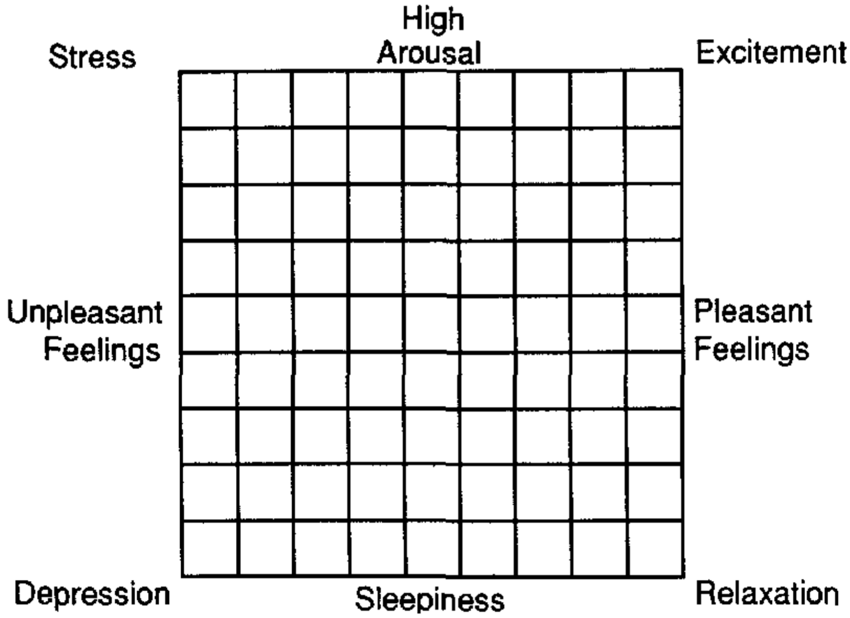

Figure 2.1: Affect grid [152].

2.1.2 Emotion

An emotion is defined by Ekkekakis [40] as a short, intense episode which is often triggered by or associated with a real or imagined person, event or thing in the past, present or future. Due to the short-term nature, one can experience a wide range of emotions within a short period of time which can influence one's overall affect. Ekman [41] argues that the majority of emotional episodes only last a few seconds each, and longer episodes lasting minutes or more are likely to be caused by repeated evocation. Emotion classification has been a topic of discussion for

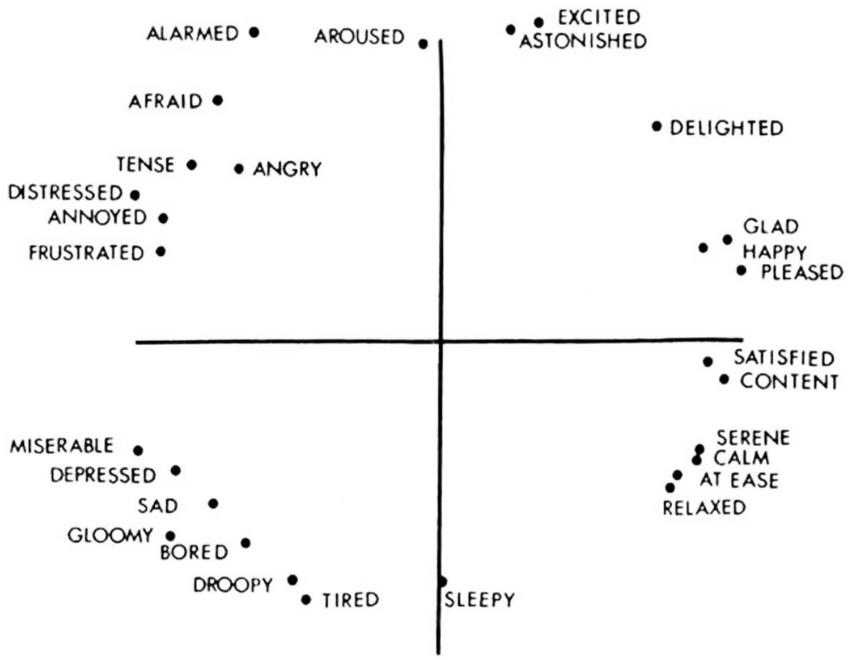

Figure 2.2: Circumplex model of affect [151].

many years and is still under debate, with psychologists attempting to define a set of basic, primary emotions from which all others stem. Following research on facial expressions, Ekman et al. [42] found evidence for six basic emotions: happiness, surprise, fear, sadness, anger and disgust. However, cultural considerations must be taken into account as some cultures may not distinguish between particular emotions both verbally and facially. Later, Plutchik [137] developed the wheel of emotions which contains eight primary emotions grouped into positive and negative polar opposites, suggesting that terms such as joy and sadness and interest and distraction are inversely related (Figure 2.3). The wheel of emotions also helps to visualise that emotions not only have a distinct classification but that they also vary on a “continuum of intensity” [54] (e.g. rage is a more intense state of annoyance). The research behind discovering the set of basic emotions is ongoing and the result of which will help to provide a succinct set of classifications for researchers to use when assessing emotional state.

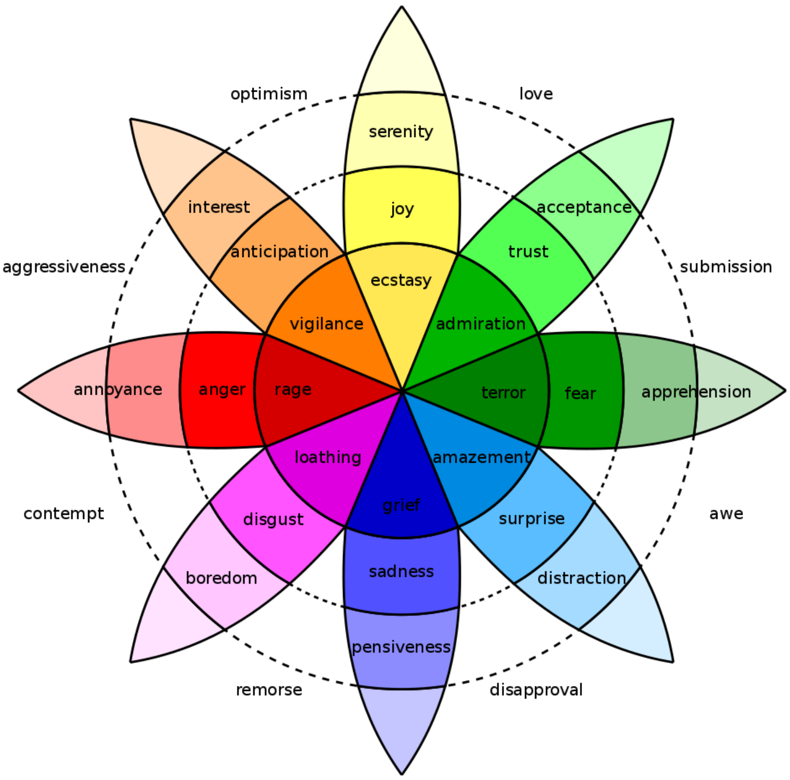

Figure 2.3: Plutchik's wheel of emotions [137].

2.1.3 Mood

Mood, in comparison with emotion, is a longer-lasting, less intense psychological state which is not usually triggered by an obvious, immediate stimulus. For example, one can simply wake up in a good mood. Unlike emotional stimuli, which often immediately precede the emotional response, mood stimuli are usually temporally distant from the response making the cause difficult to identify [116]. As with emotional episodes, the specific time scale in which mood persists is not fully understood; however, some researchers have evidence to suggest that a specific mood state can continue across multiple days [97] and should last for at least several minutes in order to distinguish it from an emotional response [108].

2.1.4 Affective Phenomena

The definitions of affect, mood and emotion, collectively known as affective phenomena, may still be subject to different interpretations; however, their distinctions and connections are clear. Affect is experienced constantly with varying levels of intensity. Emotions are short, intense episodes with clear triggers and moods are longer and less intense, sometimes without an obvious trigger. Affect can be a component of, and influenced by, emotion and mood. In the example by Russell [150], pride is feeling good about oneself. "Feeling good" is the affect and "pride" is the emotion. Mood and emotion can also influence one another. For example, experiencing several positive emotions may help to put us in a good mood; however, someone in a bad mood may be more likely to experience negative emotions. These three affective phenomena are experienced collectively to form our emotional state.

2.1.5 Wellness & Well-Being

Psychological wellness, or well-being, is a term to describe a person's general condition in addition to their mental health. Several models including Diener's tripartite model of subjective well-being (SWB) [35] and Ryff's six-factor model of psychological well-being (PWB) [155] aim to define distinct components which contribute to a person's wellness. Diener's model considers frequent positive affect, infrequent negative affect and high satisfaction with life as key components of wellness. Ryff's model considers six distinct dimensions of wellness including autonomy, environmental mastery, personal growth, positive relations with others, purpose in life and self-acceptance. The two models describe components which are very different from one another but both measure aspects of a positive life. Flourishing, coined by Keyes [71], is another term often used by psychologists to describe a person who is living "within an optimal range of human functioning, one that connotes goodness, generativity, growth and resilience" [49]. Further components of wellness include identity, self-worth, self-control, self-care and emotional awareness. As with the classification of emotion, no single model can adequately and concisely incorporate all aspects of well-being, rather all of the aforementioned components unify to describe one person's psychological wellness.

2.1.6 Mental Conditions

Depression, anxiety, bipolar disorder, social withdrawal and suicidal thoughts are just a few of the many mental conditions which affect people worldwide. Although this thesis will not focus on tracking, detection or interventions for specific mental conditions, the work will make contributions to the detection and interventions of core psychological states which are components of those mental conditions. For example, high negative affect and low positive affect are contributors to a diagnosis of depression [195] and being able to detect a sudden drop in affective state may be used to predict an oncoming depressive episode.

2.2 Psychological Assessment

Recording psychological phenomena is one of the first steps to understanding and monitoring how a person feels. These assessments are fundamental to the majority of research in the psychology domain and have been developed and revised to record many different aspects of a person's psychological state. In order to obtain the ground truth of a person's psychological state, researchers can either rely on self-reports from the participant or observations of the participant. This section focuses on self-reports as the method of data collection for the core affective phenomena discussed in Section 2.1.

2.2.1 Experience Sampling Methodologies

The method in which self-reports are obtained is its own area of research and the design and limitations must be considered when conducting a study. Collectively known as "experience sampling", there are many variations to this data collection methodology. Ecological momentary assessment (EMA), diary methods and the day reconstruction method are some of the most prominent in the literature.

"Ecological momentary assessment involves repeated sampling of subjects' current behaviors and experiences in real time, in subjects' natural environments" [167]. The seminal papers by Shiffman et al. [167] and Bolger et al. [15] aim to minimise recall bias and maximise ecological validity by providing detailed guidance on all aspects of momentary assessment. Both papers on EMA and diary methods discuss three types of experimental design: event-based, time-based and combination designs.

Research cases where specific events dictate the moment when data should be collected should follow an event-based design. For example, a participant might be asked to manually report their emotional state immediately after a social interaction. A limitation of event-based design is that the knowledge of an impending self-report may cause the participant to alter their behaviour in order to avoid the assessment interruption. The frequency of self-report triggers is therefore very important to maximise compliance and can be remedied by carefully considering the frequency of events and whether a random or combined sample of those events would be sufficient for analysis. Falsification is also another constraint to consider, where participants may not report an event which did occur or may falsely report an event which did not occur. This could be solved through the automatic detection of such events using technology; however, this solution poses its own challenges which are discussed in Section 2.3.

A time-based design involves participants completing self-reports at temporal intervals, for example monitoring mood changes over multiple days, rather than being triggered by a specific event. Participants are reminded to complete their assessments via some method of signalling such as an alarm or smartphone notification. Similar compliance issues may arise with equally spaced intervals, thus some designs opt to use non-specific intervals (e.g. "end of the day", which will have differing interpretations) or randomised times within a set of fixed intervals, known as stratified random sampling. Assessment times are also sometimes limited to waking hours depending on the goal of the study [167].

By requiring participants to complete self-reports at the moment an event occurs or at various times throughout the day, the ecological validity can be maximised. This immediacy of the assessments also helps to reduce recall bias which can occur if participants are asked to submit self-reports in hindsight. Combining event-based and time-based designs can be useful in situations where researchers want to understand how assessments during a specific event compare with base-rates. For example, a study may want to collect data about how participants feel during a panic attack in comparison with their average psychological state, thus collecting data when the participant has an attack is equally as important as collecting frequent data during the rest of their day.

The Day Reconstruction Method (DRM) by Kahneman et al. [70] is "a survey method for characterizing daily life experiences" by assessing how people experience activities throughout their day. The DRM boasts the ability to recover a sequence of affective experiences from the previous day without significantly increasing recall bias and participant burden. This method also reduces frequent disruption to daily activities and ultimately provides a dataset containing a full sequence of events and assessments rather than a set of individual momentary samples from the day. Although the sequential nature of the activity recall can help to reduce recall bias, the participant burden may still remain high depending on the length of the specific assessments required from the participants. For example, if participants are asked to answer a single question on a 5-point Likert scale every hour, this may be considered less burdensome than having to recall a day's worth of events and their associated emotional impact in one session. On the contrary, a single session can be planned for in advance unlike randomised interruptions. This method of assessment may also not be appropriate for participants with memory conditions such as Alzheimer's disease who may not be able to accurately recall prior events. The asymmetric psychological effect of positive and negative occurrences holding a higher precedence in memory [178, 190] may also affect recall bias regardless of the sequencing methodology.

2.2.2 Traditional Psychological Assessments

An important step in studying changes in emotional well-being is to ensure that the correct assessment tools are chosen in order to answer the proposed research questions. There are a vast selection of psychological assessments used for specific aspects of a participant's psychological state and this section will cover those which are most common when assessing affect, emotion, mood and well-being.

2.2.2.1 Assessing Affect

Arguably the most prevalent survey of affect is the Positive and Negative Affect Schedule (PANAS) [196]. Originally developed in 1988 by Watson et al. [196], the survey consists of a positive and negative scale each containing 10 items such as excited, determined, distressed and nervous. The PANAS demonstrates largely uncorrelated scales which are stable at a wide range of time periods (from momentary to over two months). The survey has since been revised in 1999 by Watson and Clark [193] with an expanded version called PANAS-X containing 60 items including additional affective states and positive and negative emotional states. The International PANAS Short Form (I-PANAS-SF) was later developed by Thompson [181] in 2007 with the goal of reducing the number of items to 10 whilst also ensuring that the emotional words used for the items are cross-culturally understandable to those whose native language is not English.

In contrast to the PANAS which measures affect across positive and negative scales, the affect grid (Figure 2.1) can also be used as a quick way to measure affect across the valence–arousal dimensions. The participant can mark their affective state within the affect grid which is divided into four quadrants representing positive high energy (top-right), positive low energy (bottom-right), negative high energy (top-left) and negative low energy (bottom-left). The further the mark is placed from the centre of the grid, the stronger the affective state. It has been argued that these two different scales (positive–negative and valence–arousal) for assessing affect are interrelated [108]. The positive and negative affect scales exist in the same factor space as valence and arousal, but are rotated approximately 45 degrees meaning that high positive affect is associated with high valence and arousal and vice versa (Figure 2.4).

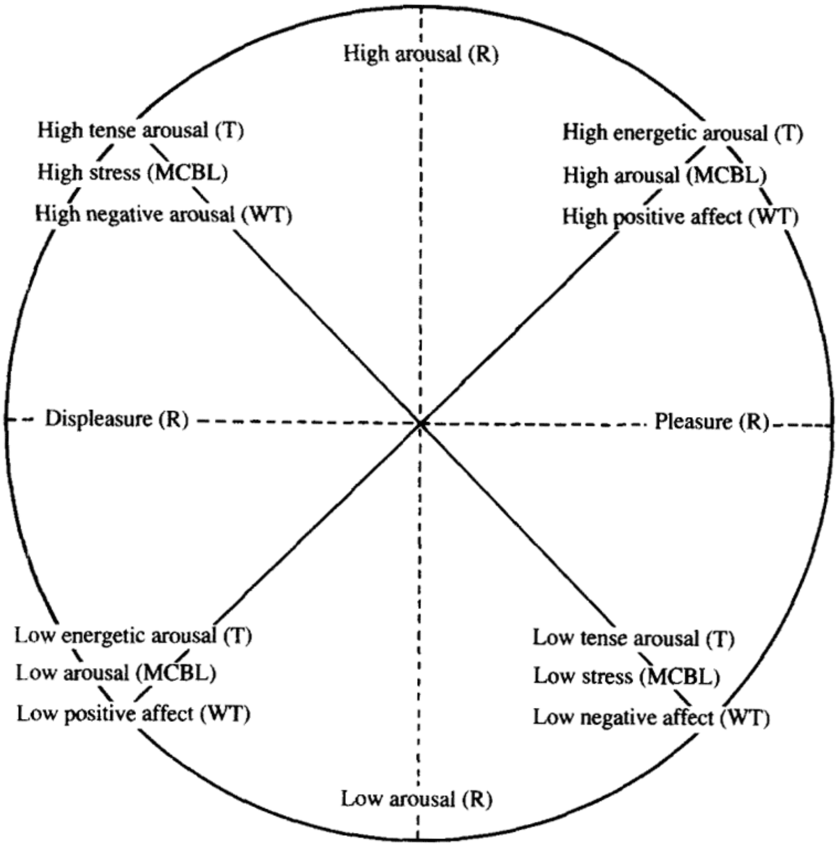

Figure 2.4: Alignment of the two measures of affect by Matthews et al. [108]. Key: MCBL [100], R [149], T [179], WT [194].

The Self-Assessment Manikin (SAM) [17] and Affective Slider [13] are non-orthogonal variations of the affect grid. "SAM is a non-verbal pictorial assessment technique that directly measures the pleasure, arousal, and dominance associated with a person's affective reaction to a wide variety of stimuli" [17]. Participants are required to select one of five items from three pictorial scales representing valence, arousal and dominance (Figure 2.5). SAM boasts a short completion time and cross-language validity as no written labels or items are used. Although the visual aspect of SAM has been praised for its age-, culture- and language-free design and causing less fatigue for repeated measures, it is not without limitations [114]. Due to the difficulty in visually representing affective states, especially arousal and dominance, the dimensions are not immediately clear without some level of verbal or written explanation. The original study assessing the feasibility of SAM included a list of words at the ends of each scale to help the participants understand what the scales represent [17]. AniSAM/AniAvatar [170] is a revision of SAM which updated the scales with more realistic looking avatars which focus more on the facial expressions to improve recognition of valence and an animated beating heart and vibrating avatar to better represent arousal (Figure 2.6). Results indicated that both valence and arousal were reported more accurately compared with the original scale, however AniSAM/AniAvatar is limited to mediums which are able to display animated visuals.

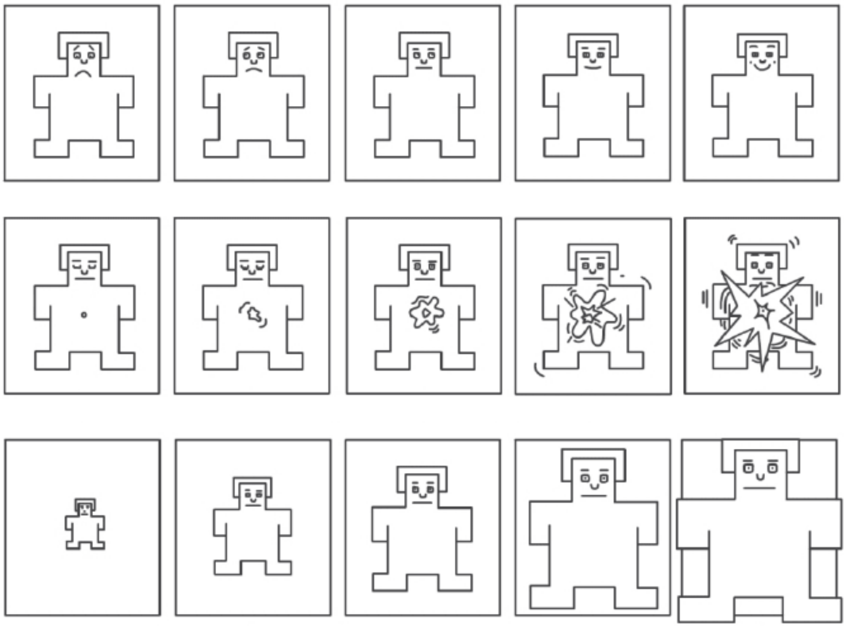

Figure 2.5: The Self-Assessment Manikin (SAM) used to rate the affective dimensions of valence (top row), arousal (middle row) and dominance (bottom row) [17].

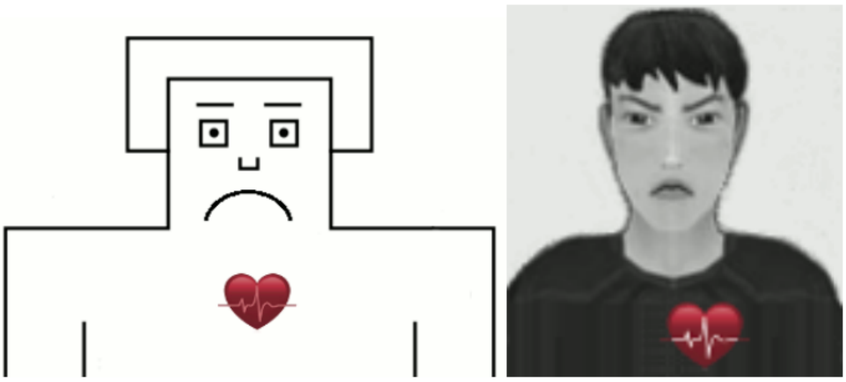

Figure 2.6: Examples of new avatars used in the AniSAM and AniAvatar scales [170].

The Affective Slider [13] is a digital update to SAM focused on modern design principles which comprises of two continuous slider controls for valence and arousal (Figure 2.7). The study uses more commonly understood emoticons at the ends of each slider to represent sleepy and awake (top) and unhappy and happy (bottom). Although these emoticons are an improvement in comparison with the avatars used in SAM, they still remain slightly unclear. The sleepy emoticon might be misconstrued as disappointment and the awake emoticon as frightened. Regardless of the potential limitations, the Affective Slider has been validated as equivalent to SAM in the assessment of valence and arousal without the need for written instructions.

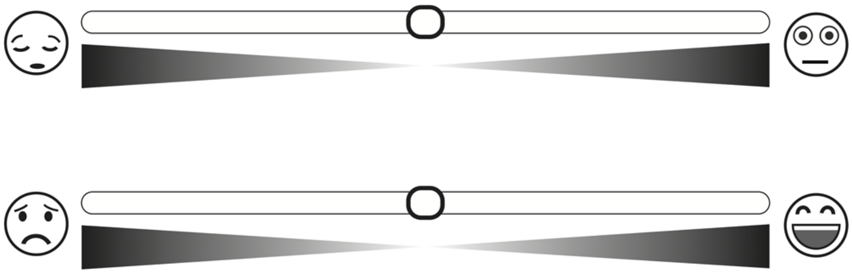

Figure 2.7: The Affective Sliders for valence (bottom) and arousal (top) [13].

2.2.2.2 Assessing Emotion

Emotional states are often used in self-assessments in order to determine a more general affective state like in PANAS or to diagnose a specific condition using surveys like Beck's Anxiety and Depression Inventories [10, 11] which assess emotions including sadness, guilt and fear alongside other factors of anxiety and depression. These emotions are usually assessed via some form of Likert scale, either with numerical values or written descriptions representing different levels of intensity. Since emotional states are usually associated with a stimulus, it is difficult to assess emotions in general without a source. An exception could be an emotion like happiness which could be more easily generalised; however, this would arguably be an assessment of one's positive affect or mood rather than the emotion itself if it is not associated with a stimulus. Researchers therefore use emotional scales to assess their participant's reactions to a particular stimulus that they are studying. Izard's Differential Emotion Scale (DES-IV) [67] utilises a 5-point Likert scale ranging from "rarely/never" to "very often" to assess 12 subscales including: interest, joy, surprise, sadness, anger, disgust, contempt, fear, guilt, shame, shyness and self-hostility. Many individual emotional scales have also been developed to thoroughly assess the intensity of said emotion in response to a stimulus. Observing the literature for just some of the numerous emotional states, happiness can be assessed using the Subjective Happiness Scale (SHS) [96] and the Authentic Happiness Inventory (AHI) [129], anger with the State-Trait Anger Scale (STAS) [172], fear using the Fear Survey Schedule-II (FSS-II) [50] and guilt and shame with the Test of Self-Conscious Affect (TOSCA) [177]. In contrast to the DES, the five aforementioned individual emotional scales rely on the participant rating their response to specific statements such as, "I consider myself a very happy person" (from SHS) which are then combined to produce a final score for that emotion.

As suggested by Scherer [162], a problem with assessing emotions in a such a manner is that of priming the participant, i.e. the participant may respond to the options in a way that they may not have done if the options were not present. Scherer [162] suggests that by using his "free response measurement of emotional feeling – the Geneva Affect Label Coder (GALC)", one can alleviate the issue of priming. The GALC functions by taking a free-form emotional written description by the participant and using a lookup table to search for synonyms or words related with a specific emotion. Using this method, an emotion can be derived without the need to present the participant with a set of predefined options.

2.2.2.3 Assessing Mood

Assessing mood as a single scale has less prominence in the literature, as assessments of affect are more commonly used; however, mood can still be easily assessed independently. Mood is commonly described as ranging from "good/positive" to "bad/negative", thus it is common to use a bipolar scale with some form of "positive" and "negative" labelling. This type of scale is often used by more modern, smartphone mood logging systems (discussed further in Section 2.2.3) due to its ability to be quickly completed.

For the sake of completeness, the Profile of Mood States (POMS) [111] should be discussed to avoid confusion between its name and its assessments. Contrary to the measure's title, POMS is a measure of psychological distress rather than that of general mood. The assessment is conducted with 65 items (or 37 in the short form) resulting in 7 subscales including tension–anxiety, depression–dejection, anger–hostility, vigour–activity, fatigue–inertia, confusion–bewilderment and total mood disturbance. These results are a blend of mood states, emotions and mental conditions and should therefore be approached with caution when attempting to assess one of the affective phenomena individually.

2.2.2.4 Assessing Wellness & Well-Being

There are also a variety of measures which can assess other aspects of wellness and well-being discussed in Section 2.1.5. Diener's tripartite model of subjective well-being [35] combines results from PANAS [196] and the Satisfaction with Life Scale (SWLS) [36] in order to produce a subjective measure of quality of life. An interesting aspect of this model is how it combines measures across different time scales: the momentary affective states and the longer lasting state of satisfaction with life. Compared with the 5-item scale of the SWLS, the Life Satisfaction Index [47] is a longer-form assessment of quality of life containing 20 items (or 11 items in the short form). Ryff's Psychological Well-Being scales are used to directly assess his six-factor model [155], Diener's Flourishing Scale [37] can assess human flourishing, Ryan & Frederick's Subjective Vitality Scale [154] focuses on the feeling of being alive and alert and the Measurement of Self-Actualization by Jones and Crandall [69] is used to assess one's feelings of potential and fulfilment.

2.2.3 Digital Psychological Assessments

Many of the traditional psychological assessments were developed before computers and smartphones became widespread and were completed by hand, either by the researchers or by the participants themselves. The proliferation of digital devices has caused a rise in psychological experiments being conducted digitally and remotely over the internet rather than in-person with a "pen and paper". What follows is a showcase of current literature which makes use of digital ecological momentary assessments in addition to some of the most popular well-being smartphone applications in distribution.

2.2.3.1 Assessing Affect

The affect grid is widely used in smartphone applications due to its ease of completion. It is used in a web application called MoodMap by Fessl et al. [46] to capture the affect of people during virtual meetings and a smartphone application of the same name to study emotional self-awareness [115]. It should be noted that these papers used colour in their affect grid designs (Figures 2.8, 2.10). Fessl et al. [46] specifically mention that the colour wheel design is based on Itten's colour system [66] (related with colour contrast) and Morris et al. [115] do not provide any reasoning for their choice of colours. It has been shown that different colours can influence mood [79] and can be interpreted differently across cultures [68], thus keeping the colour of the affect grid neutral may be advised to remove any unwanted side effects.

Some researchers have opted to use non-orthogonal versions of affect scales because range sliders are a common user interface (UI) design pattern for web and smartphone interfaces. One paper conducted assessments of affect using a 6-item scale of range sliders resulting in a three-dimensional model comprising of valence (positive–negative), calmness (restless–relaxed) and arousal (tired–awake) [44]. Additionally, AffectAura [109] attempted to develop a new way of visualising three dimensions of affect (valence, arousal and engagement) akin to the Self-Assessment Manikin [17]. The interface displays affective states on a timeline, with each state represented as a "bubble" (Figure 2.9). The colour represents valence, the "burst" shape represents arousal (similar to SAM) and the opacity represents engagement. The size and height of the bubbles were used to represent other aspects of the study. This visualisation suffers from the same limitations as SAM in that the scales are not immediately obvious without a description. Lastly, in contrast to the affect grids and sliders, there were also studies [192] which chose to use the Photographic Affect Meter (PAM) [138] that required participants to indicate how they felt from a selection of images which are mapped to the valence and arousal affect scales (Figure 2.11). This method has the potential of being easier to understand than affect grids or scales with confusing or ambiguous labels.

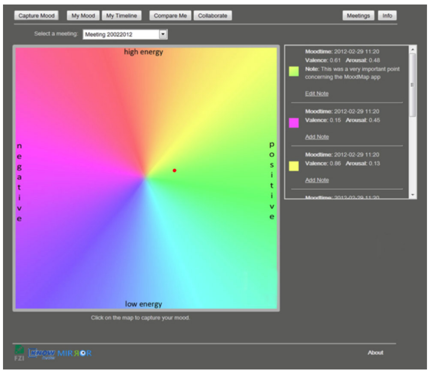

Figure 2.8: An affect grid to capture the affect of people during virtual meetings [46].

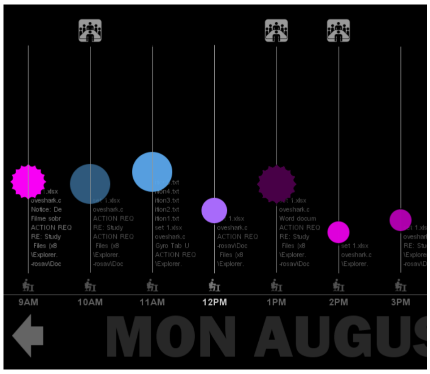

Figure 2.9: The AffectAura interface, representing affect using various visual aspects of "bubbles" [109].

Figure 2.10: Using an affect grid on an old Android device as part of a study researching emotional self-awareness [115].

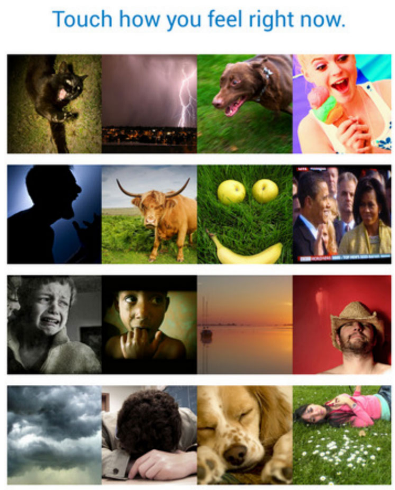

Figure 2.11: The Photographic Affect Meter (PAM) [138] that requires participants to indicate how they feel from a selection of images which are mapped to the valence and arousal affect scales [192].

2.2.3.2 Assessing Emotion

Emotional states in applications are mostly assessed using different forms of Likert scales. Sliding scales and multiple choice answers can be used to quickly assess the state of multiple emotions and conditions (Figure 2.12). These studies involved a variety of cases such as emotional self-awareness [115], assessing mental health and academic performance of college students [192], mood recognition at work [200] and a borderline personality disorder and substance use disorder smartphone therapy [147]. The T2 Mood Tracker app [32] can be used to track a variety of emotional experiences including general well-being, anxiety and depression and can also be customised to address any issue in need of assessment. Single emotional scales such as happiness can also be assessed very quickly as shown in the studies analysing how happiness changes depending on time and activities [156] and in different locations across the UK using an app called Mappiness [101].

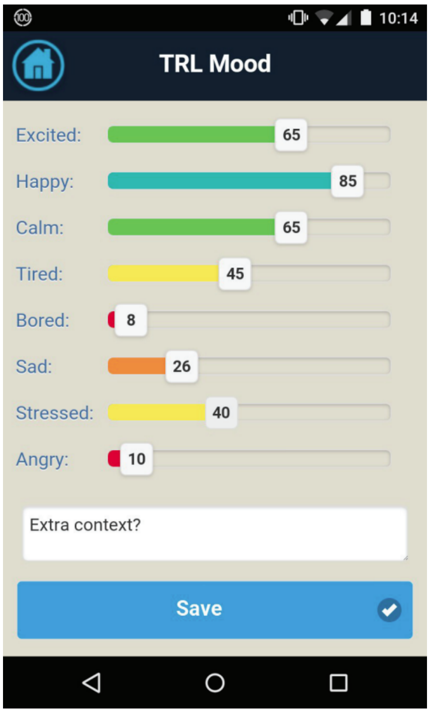

Figure 2.12: The HealthyOffice Android app allowing users to self-report various emotional states using sliders. [200].

Figure 2.13: The Peas Android app allowing users to self-report their emotional state by selecting an emoji [117].

Aside from Likert scales, non-academic applications (those with no research foundation) such as Mood Runner [64], Feelic [106] and Peas [117] (e.g. Peas, Figure 2.13) use emojis to identify different "feelings" that the user may wish to record. The assessment method used in these examples is limited as users can only choose a single feeling at a time and there is no way to declare the intensity of the chosen feeling (e.g. very happy or slightly happy).

2.2.3.3 Assessing Mood

As discussed in Section 2.2.2, mood is widely used in digital psychological assessments due to its quick rate of completion. Many non-academic smartphone applications use a 5-point Likert scale augmented with emojis, and sometimes labels, to symbolise the positive to negative scale of mood [112, 183, 186, 188, 197, 201] (e.g. Daylio, Figure 2.14). Other apps do not use emojis and instead implement a traditional Likert scale ranging from 3 to 10 points, either with numeric or text labels [14, 55, 63, 113, 122, 182] (e.g. iMoodJournal, Figure 2.15). In general these apps tend to provide digital diary-style experiences, allowing users to log their daily mood

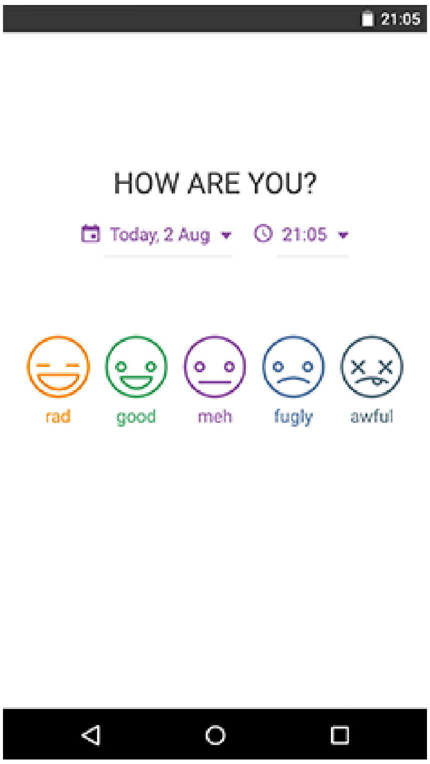

Figure 2.14: The Daylio Android app allows users to report their mood on a 5-point emoji Likert scale [197].

Figure 2.15: The iMoodJournal Android app allows users to report their mood on a 10-point textual Likert scale [63].

and add extra tags and notes for more detail. They also provide visualisations in the form of graphs, calendars and timelines to easily observe mood changes over time.

A major concern with app- or web-based mood/emotion trackers is that many are created by amateur developers and enthusiasts without any medical or psychological background or research. Mental health is a sensitive subject and non-academic applications that have the potential to reach thousands of users can have a major negative impact on their users if not built correctly. Developers should ensure that they are designing their apps with research guidelines in mind and have sought guidance from medical professionals.

2.2.4 Challenges

There are many challenges and concerns with the traditional and digital assessment methodologies which can lead to inaccurate, or loss of, data due to falsification, compliance issues, satisficing and misunderstanding.

Trust is placed in the participants to complete self-reports accurately and truthfully. Falsification of data is a concern especially with data involving sensitive subjects such as mental health; participants may be worried or embarrassed to share their weaknesses and vulnerabilities with researchers. The participant's trust and comfort with the researchers and the study is important to ensure that data is accurate.

Compliance and drop-out rates are further concerns for researchers that must be considered during the design of the study methodology. Studies which do not take place in-person, and which have a long duration, are especially vulnerable to compliance issues as participants experience fatigue with the study. The immediacy of smartphone applications means that short-form surveys may be more appropriate and might be why methods like the affect grid and emotional sliders are often used. Using automated event-triggers coupled with a recommendation system may also be a potential way to improve compliance rates by ensuring that it is an opportune moment for the participant to complete a survey [134]. Event-based designs also suffer from limitations. Participants must fully understand the definition of the trigger, which can be problematic in certain scenarios: "If subjects are to make a record every time they eat, does chewing gum count as eating?" [167]. The effect of the assessment itself must also be taken into consideration, such that if the purpose of the study is to monitor stress, but the self-reporting methodology causes more stress to the participant, the results may become invalid. This particular phenomena is discussed further in Section 2.4 where it can be used to the researcher's advantage via the *intervention bias*.

It was proposed that, when answering survey questions which would require substantial cognitive effort, some participants will provide an answer that requires the minimum amount of thought, known as satisficing [78]. For example, a participant may log the same mood state every day, perhaps because that value was selected by default or because it was the easiest to choose from the interface. Krosnick [78] discussed potential methods to avoid satisficing including randomising the order of response options, which is especially easy to accomplish for surveys with multiple subscales. For example, in the PANAS [196], the words for both the positive and negative scales are shuffled rather than being grouped together. For surveys containing written statements that must be agreed or disagreed with, the linguistic polarity of some of the statements are often reversed to ensure that the participant applies more thought to their answers.

The wording and descriptions used for psychological surveys is another important factor for researchers to consider. Many common psychological surveys have international or child-friendly alternatives with simpler language (e.g. PANAS and I-PANAS-SF). For self-reports involving recall, it is important that the time period is made clear, for example: "How do you rate your mood right now?" vs. "How was your mood today?". Lastly, one should take care when using terms which may be clear to psychologists and academics but may be misconstrued by non-academics taking part in research studies or using self-help apps. For example, the term "arousal" may be mistaken for sexual arousal rather than alertness or attentiveness. This has been found to be the case with the affect grid where researchers tend to use different labels than the original design [152]. Fessl et al. [46] and Morris et al. [115] used "energy" and "positive–negative", Lathia [82] used "alert–sleepy" and "positive–negative" and the original by Russell et al. [152] used "high arousal–sleepiness" and "pleasant and unpleasant feelings". This raises problems because the labels are not congruent; a low-arousal state of calmness is not synonymous with sleepiness. Further work is required in order to provide a consistent standard for these labels to ensure that academics and non-academics alike can derive the same meaning from them.

2.3 Passive Detection of Emotional Well-Being

Although assessments of emotional well-being via self-reports have been prevalent for many years, it comes with an inherent drawback: it requires active contribution from the participant. This introduces the issues discussed in Section 2.2.4 such as burden, falsification, compliance, satisficing and linguistic problems. The advancements in technology have made it possible to passively detect different aspects of emotional well-being without any, or minimal, user input. For studies which would normally require ecological momentary assessments, this can greatly reduce the burden on the participants and solve some of the assessment challenges. This section aims to outline the research using smartphone and wearable sensing and online behaviour in order to passively detect different aspects of emotional well-being.

2.3.1 Smartphone Sensing

The majority of smartphone devices contain a multitude of sensors which can be used for scientific research. These sensors include the accelerometer, compass, GPS, gyroscope and proximity as well as the touch screen, microphone and front and back cameras. Researchers have discovered many innovative and inventive ways to use these sensors for assessing different aspects of emotional well-being.

Location is a common feature used in this field to analyse mobility patterns as they have been closely related to physical and social activity and emotional well-being. The GPS sensor in smartphones can detect a user's precise location with high levels of accuracy and can collect data passively in the background. A study by Wang et al. [192] used location data to detect mobility patterns and found a correlation between the distance travelled and loneliness. Sandstrom et al. [158] published a paper analysing how location can have an impact on one's affective state, concluding that users reported more positive affect in social locations, such as a family or friend's house, or at a restaurant, café or pub, rather than at home. Canzian and Musolesi [19] also analysed mobility patterns from GPS location data in order to predict depressive mood states.

Many researchers combine location with other sensors and behavioural features that can be passively tracked in order to infer emotional well-being. Wang et al. [192] used data from the accelerometer, proximity sensor and microphone to assess the day-to-day impact of the student workload on affect, stress, sleep, physical activity, sociability and academic performance. Results in a paper by Lee et al. [83] show that their system can classify seven emotions (happiness, surprise, anger, disgust, sadness, fear and neutral) using motion, location and ambient light sensors alongside typing behaviour features via a Twitter client. A similar system by Ma et al. [98] can classify displeasure, tiredness and tensity dimensions using the accelerometer, GPS, proximity and microphone alongside SMS and call logs. BeWell, a smartphone application by Lane et al. [81], monitors the user's sleep, exercise and social interaction using the GPS, accelerometer and microphone to produce three well-being scores for each behaviour which are used for an intervention. Finally, MoodScope by Li Kam Wa et al. [87] uses location traces and smartphone usage (app usage, calls, SMS, email and web browsing) to estimate the user's valence and arousal.

Additional work in this field uses sensors such as the microphone and camera for their inference systems. Rachuri et al. [145] used voice samples recorded from the microphone to detect individual emotions including happiness, sadness, fear, anger and neutral. The platform can also infer physical activity from the accelerometer and perform speaker recognition and co-location using both the microphone and Bluetooth to detect proximity. A study by McDuff et al. [110] from the emotion measurement technology company Affectiva demonstrated the use of the smartphone camera to recognise emotional facial expressions in real-time including anger, disgust, fear, joy, sadness, surprise and contempt.

Many, but not all, of these papers combine sensor data with activity and behavioural data to produce their predictions with varying levels of accuracy. This may suggest that sensing, activity or behavioural data alone is less sufficient to accurately predict emotional well-being and that a combination of features may produce more fruitful results.

2.3.2 Wearable Sensing

In addition to smartphone sensing, using sensors in wearable devices to passively detect various aspects of emotional well-being is also prevalent in the literature. Wearable devices such as smartwatches, wristbands and headsets can contain a variety of sensors including an accelerometer, gyroscope, proximity and microphone as well as sensors for measuring physiological signals such as heart rate, galvanic skin response (GSR), skin temperature, electrocardiogram (ECG, heart) and electroencephalogram (EEG, brain). The following research utilises data from different variations of wearable sensors to train classifiers to automatically detect different psychological states including emotion, mood and stress. The classes were labelled using self-reported ground truth or specific elicited emotions depending on the study design.

A paper by Picard et al. [133] from 2001 used blood pressure, skin conductance, respiration and electromyogram (EMG, muscles) sensors to predict eight emotional states: anger, hate, grief, platonic love, romantic love, joy, reverence and neutral. Although these sensors were attached to a wearable computer unit, it would not be considered as convenient as more modern wearables such as a fitness wristband. Similarly Haag et al. [57] were able to classify affective valence and arousal using a bulky setup of EMG, ECG, respiratory, skin conductance, skin temperature and pulse rate sensors. In comparison, Zenonos et al. [200] used a single wearable chest sensor to predict eight different emotional states (excitement, happiness, calmness, tiredness, boredom, sadness, stress and anger) by collecting heart rate, pulse rate, skin temperature and accelerometer data. Lisetti and Nasoz [90] used GSR, skin temperature, accelerometer, heat-flux and near-body ambient temperature sensors in a wearable armband to predict six emotional states: sadness, anger, fear, surprise, frustration and amusement. Sourina and Liu [171] used the Emotiv EEG headset to propose an arousal–valence recognition model to passively detect the user's psychological state regardless of any visible physical behaviours. Exler et al. [44] used heart rate data from a smartwatch alongside smartphone sensor data to detect valence, arousal and calmness states. Sas et al. [161] used the GSR sensor in a SenseWear Pro_2 armband to identify peaks of emotional arousal in order to filter emotionally important photos captured with a wearable camera. Sano and Picard [159] used accelerometer and skin conductance sensors in a wristband combined with smartphone sensors and activity to classify levels of stress. de Santos Sierra et al. [31] also conducted some valuable research on stress detection, concluding that using only heart rate and galvanic skin response data can be sufficient for accurate detection. This is important because these are common sensors found in smaller wearables such as fitness wristbands, thus promoting the idea of less intrusive passive detection. Additionally, Setz et al. [164] used only skin conductance to differentiate between stress and cognitive load in a working environment to ensure that it is the emotional response being captured rather than a normal level of mental activity.

2.3.3 Challenges with Smartphone and Wearable Sensing

As electronics have miniaturised over the years, smartphones and wearables have become increasingly relevant in the research for the passive detection of emotional well-being. Depending on the factor being observed, it may be required that the participant uses or wears the device for an extended period of time, thus ensuring that it is comfortable and non-intrusive is of utmost importance to ensure data collection is successful. Devices will inevitably continue to reduce in size and intrusiveness; however, one of the limitations at the current time is battery life, especially when running apps with continuous sensing. Without an advancement in battery technology, wearables will remain inconvenient to users who have to regularly charge the devices which may not be collecting important data during that time.

Furthermore, passive sensing is often associated with considerable privacy concerns. Continuous tracking of location, audio or the camera in the background can be extremely intrusive regardless of the researcher's trustworthiness. For example, the location of a person's home can very easily be inferred using only their raw location traces [19]. Even sensors such as the accelerometer, which might not appear to be as intrusive, can be used for malicious purposes. Owusu et al. [121] were able to demonstrate that accelerometer readings are sensitive enough to be used to spy on keystrokes, such as when typing a password on a smartphone.

The demographics of the target audience must also be considered with these types of technologies. The age and mental capabilities of the participants may influence their ability to properly interact with the technologies or follow the study guidelines. The availability of the technologies may also be an issue in certain demographics. For an online application where the participants need to use their own smartphone, this may not be appropriate for certain ages or for particular countries or locations where smartphones and high-speed internet are not as widespread. As the proliferation of smartphone, wearable and internet technologies become more universal, and the population becomes more technology-literate, these demographical challenges will diminish.

2.3.4 Online Behaviour

Tracking behavioural patterns can reveal a great deal about a person's mental state. The popularity of online social networks has given rise to a culture in which it is normal to share aspects of one's personal life, sometimes publicly, on the internet. This grants a vast amount of data for social scientists who wish to passively observe individuals without the need for any personal interaction. Literature in this area explores the use of online data to passively track different aspects of emotional well-being. This data broadly consists of textual posts, media posts (images and videos) and online behavioural patterns.

The analysis of textual data to extract useful emotional meaning is known as sentiment analysis [124]. Understanding the methods and limitations of this process is fundamental in this field of research where the data is often online textual content. The high-level process involves inputting a string of text and outputting meaningful emotional data such as the number of emotional words. One of the most popular and commonly used tools for sentiment analysis is the Linguistic Inquiry and Word Count software (LIWC) [92] which uses large labelled dictionaries to classify words into psychological categories such as positive, negative, anxiety, anger and sad. The number of words assigned to each category can then be used to determine the sentiment of the passage as a whole. For example, extracting the number of positive and negative words can allow us to ascertain the emotional polarity of the sentence. Performing sentiment analysis on online data comes with several complications which are the subject of further research in the field. For example, unlike formally written documents, online text can often contain slang, abbreviations, misspellings, sarcasm, emojis and emoticons. Research is being conducted to be able to automatically detect and quantify occurrences of such complications [1, 73, 75, 185].

Twitter is an attractive platform for social scientists as the majority of the mostly textual data is publicly accessible. Facebook also has a wealth of data of all different mediums; however, the platform is less public and requires more stringent procedures to be granted access to the data. There is also some research into using text and image analysis on content from photography-focused platforms such as Instagram and Flickr; however, as of July 2018, Instagram's public API (Application Programming Interface) has been deprecated for non-business use [56].

The literature encompassing online behaviour can be loosely grouped into studies which use online data to model an individual's emotional well-being and studies of the general sentiment of a population.

2.3.4.1 Individualised Models

Agarwal et al. [1], Balabantaray et al. [7] and Pak and Paroubek [123] show how positive, negative and neutral sentiments and individual emotions can be classified from individual tweets with good accuracy (up to $\sim$73%) when compared with manual annotations as ground-truth. Once extracted, this low-level classification can be used to determine the positivity or negativity of that individual's sentiment and may be used as a feature for other aspects of emotional well-being. As proposed by Roshanaei et al. [148], additional features to supplement sentiment analysis can be extracted such as the topic of the tweet, the temporal trends of the user's posting behaviour and the demographics of the user. De Choudhury et al. [28] used the sentiment, textual content and temporal aspects of tweets in addition to the social interaction with other users in order to predict major depression in individuals ahead of their reported onset. A paper by Matsumoto and Hoashi [107], citing our earlier published work presented in Chapter 3, explored a similar approach to ours by combining data from multiple social networks. By incorporating textual tweet data with social and physical data from tweets without text (such as mentioning a friend alongside a photo using the @ symbol or sharing content from another social network such as a location from Foursquare) the researchers found significant correlations between those features and the mood of the upcoming day. Ortigosa et al. [119] utilised status updates from Facebook to perform sentiment analysis to extract the user's positive, negative or neutral sentiment in addition to detecting significant emotional changes by calculating their average sentiment and observing any abnormal deviations. Post-partum depression has been predicted using online activity, social interactions and linguistic style from Facebook [29] and De Choudhury et al. [30] attempted to predict transitions from mental health discussions to suicidal propensity of users of Reddit communities using linguistic style and social interactions. Data from Instagram has also been used as predictors for depression. Reece and Danforth [146] used hue, saturation, brightness and the number of faces from image analysis and photo filters alongside comments, likes and daily engagement as predictive markers of depression. Lup et al. [93] also show a marginal positive correlation between the duration of daily Instagram use and depressive symptoms.

2.3.4.2 Population Models

Online data can also be used to model the general well-being of a population rather than an individual. Bollen et al. [16] performed sentiment analysis on all tweets posted to Twitter for six months to observe the impact that social, political, cultural and economic events have on the population's emotional state as portrayed in their tweets. Hasan et al. [59] achieved a similar result through observing emotional tweets related to specific positive and negative public events. Gross national happiness has also been the subject of some research in this field. The "happiness" of large populations has been tracked using data from Facebook [76] and Twitter [143]. Kramer [76] compared sentiment analysis results of Facebook status updates with Diener's satisfaction with life survey. Quercia et al. [143] used sentiment analysis results of tweets to find significant correlations with that community's socio-economic well-being and Prata et al. [139] used sentiment analysis of tweets to generate mood maps across the Brazilian territories. Emotional reactions to specific locations, as opposed to a population's well-being *within* a location, can also be extracted from online data.

Quercia et al. [144] used crowdsourced emotional responses to Google Street View images to determine the types of locations people associate with beauty, quiet and happiness and Hauthal and Burghardt [60] used sentiment extracted from titles, tags and descriptions of geotagged photos on Flickr and Panoramio to associate emotion with location. In a similar manner in which companies value public opinion and engagement towards their products, population-wide emotional data can be useful for analysing the impact of important public events or for urban planning for example.

2.3.5 Challenges with Online Behaviour

Using online data for the purpose of passively tracking emotional well-being is not without limitations. Internet users are not obliged to speak or behave in the same way in which they would in real life which presents the issues of self-presentation and self-idealisation. Self-presentation is the way in which a person decides to exhibit themselves online, be it truthfully, dishonestly or somewhere in between. For example, one might not lie about their life online, but may only portray the positive aspects of their life, thus painting a false picture of their life as a whole for their social network [104]. Self-idealisation is the act of falsifying one's online presence to appear as their ideal self rather than their true self. This might involve displaying a completely different life, one in which they aspire to achieve. Manago et al. [104] suggested that OSN users, especially emerging adults, might use OSNs to "try out" ideal selves in a virtual environment before committing to them in real life once sufficient social validation is acquired. Contrarily, Back et al. [6] found that personality is an aspect which is portrayed accurately by users on Facebook, suggesting that there was no evidence of self-idealisation for those users. The decisions behind how users self-present online may be due to several factors including the type and purpose of the social network (e.g. Facebook for private posts vs. Twitter for public posts), the size of their social network [89], or simply down to user preference. This filtering of online profiles can lead to mistrust in the accuracy of what users share and how it realistically relates to their emotional well-being. A validation study by Wang et al. [191] found that Facebook's Gross National Happiness metric [76] is in fact negatively correlated with satisfaction with life, suggesting that users "may try to disguise their real emotions". Additionally, a paper by Liu et al. [91] found that negative Facebook posts were correlated with negative life satisfaction, whereas positive posts were not. A major issue found in the past literature using OSN data to predict aspects of emotional well-being was that many authors did not validate the inferred results against real-world ground truth data. Most papers would take the output at face value which does not control against these self-presentation and self-idealisation issues.

Further limitations arise from the OSN itself. Population bias (specific user demographics using specific OSNs), proprietary filtering algorithms, design for behavioural manipulation and fake accounts all present difficulties for researchers working with social networks [153]. If researchers are using OSNs to measure emotional well-being where the variable is an external factor, emotional contagion may be an unwanted side-effect of using that OSN. As in real-life, human emotions can be influenced by the emotions of others and this has been observed to occur within online social networks too through empathy [140], clusters of similarly opinionated or emotionally-congruent users (homophily) [142, 176] and emotional manipulation of news feeds [77]. Some studies have also found that the social aspects of OSNs can have negative effects on emotional well-being via negative social comparison [45, 93, 187]. Snapchat, on the other hand, is more commonly associated with positive mood, tending to be described as a more humorous and playful platform [135, 141]. Studies have shown that interactions on Snapchat, usually performed via back-and-forth sharing of photos with captions, was considered to be associated with a more positive mood in comparison with texting, email and Facebook [8].

2.4 Positive Psychology Interventions

Within the field of clinical psychology, there is a bias towards treating negative mental health issues rather than enabling, encouraging and preventing mental health disorders for those who are already flourishing.

As Seligman and Csikszentmihalyi [163] succinctly described: "psychologists have scant knowledge of what makes life worth living. They have come to understand quite a bit about how people survive and endure under conditions of adversity. However, psychologists know very little about how normal people flourish under more benign conditions. Psychology has, since World War II, become a science largely about healing. It concentrates on repairing damage within a disease model of human functioning. This almost exclusive attention to pathology neglects the fulfilled individual and the thriving community."

They go on to describe that positive psychology instead focuses on happiness, subjective well-being, satisfaction with life, human flourishing, mindfulness, positive thinking, gratitude, kindness and many other positive aspects of life. Positive psychology interventions are defined as "treatment methods or intentional activities aimed at cultivating positive feelings, positive behaviors, or positive cognitions" [169]. There are parallels between positive psychology interventions and other forms of intervention such as cognitive behavioural therapy (CBT) and mindfulness. CBT is a problem-focused and action-oriented form of therapy aimed at treating specific diagnosed mental health disorders by modifying thoughts, beliefs, attitudes and behaviours [12]. Mindfulness is the psychological process of increasing attention and awareness of the present moment and has been used to treat various psychological conditions [26]. This section explores traditional and modern positive psychology interventions to aid in the improvement of various aspects of emotional well-being.

2.4.1 Life Review Interventions

Many traditional interventions for improving emotional well-being rely on the participants expressing their experiences and emotions through writing tasks. The therapeutic effects on mental health by writing about emotional experiences has been documented in Pennebaker's 1997 paper [128] in which participants were instructed to write down their thoughts and feelings about an influential, emotional issue experienced during their lives. The paper discusses the positive effects of disclosure, especially when done so in an anonymised fashion. Another such task involves focusing on and writing about intensely positive experiences, expressing as much detail about the feelings, thoughts and emotions that were experienced during that moment [18]. A different intervention type involved group reminiscence sessions in elderly nursing homes producing a significant increase in life satisfaction by recollecting positive experiences such as holidays, school, marriage and family. News articles, music and old photos were used to stimulate topics of discussion [24]. A similar reminiscence-themed study found that individuals who took part in life review therapy showed lower levels of depression and increased life satisfaction when compared with a control group [27]. A study in Japan found correlations between kindness and subjective happiness with the happier participants showing a higher desire to be kind and a higher aptitude for recognising kindness. They additionally reported a significant increase in happiness after counting their own acts of kindness [120]. In another study, participants were required to write down five things that they were grateful for every week for ten weeks. The results displayed a significant increase in how the participants rated their lives, in addition to improved optimism and increased joy and happiness in response to being grateful about receiving aid [43]. Gratuity was the focus of another writing study where participants were instructed to think about a time when they were grateful for something that someone did for them and write a letter to that person about that experience and how it affected them [97]. A relatively less common but interesting positive intervention task involved the participant thinking about and discussing a positive thing or event in their life that might have never occurred had something in the past happened differently. In these interventions participants were invited to consider events that they may think were caused by "destiny" or "fate". A paper by Koo et al. [74] asked participants to write about such surprising events and compared them with participants who wrote about unsurprising events and found that those in the first group reported more positive feelings. In addition to writing tasks, Lichter et al. [88] found that group discussion of beliefs and attitudes about happiness and daily rehearsal of positive feeling statements showed improvements in happiness, satisfaction and depression. Furthermore, goal setting is another well documented form of positive psychology intervention. Writing about life goals has been shown to be associated with significant improvements in subjective well-being, specifically focusing on best possible future selves [72] and teaching goal-setting and planning skills [102, 165].

2.4.2 Technology-Enhanced Interventions

The widespread usage of, and technology inside, modern smartphones allows for traditional intervention methodologies to be applied digitally and remotely. A small study used controlled smartphone tasks to elicit affective states, specifically stress and relaxation [23] and another used a smartphone application called WellWave to improve psychiatric and physical well-being by promoting physical exercise using prompts, providing confidential communication with peer staff and access to motivational articles and videos [99]. A study by Howells et al. [61] compared the use of an existing mindfulness smartphone app called Headspace with a neutral control list-making app and found a significant increase in positive affect and reduced depressive symptoms in the Headspace group in comparison with the list-making group. There has also been limited research into using photos for therapeutic purposes. The majority of work focuses on assessing psychological state from the content of photos [2, 146]; however, there is not much work on how the act of photography can affect well-being. Phototherapy (not to be confused with the physical treatment involving exposure to light) is used in counselling to enhance the therapy process by using photos by, or of, the patient as a way to evoke memories and feelings [198]. Photography can function as a visual language for clients who are particularly anxious about verbalising their thoughts and emotions, by taking photos of things that represent their feelings. Reflecting upon photographs can also be a method to recognise and document development during an intervention [52]. Photography interventions outside of counselling have also had positive effects on mood, emotion, affect and life satisfaction. Participants who were asked to take three "creative, beautiful and meaningful" photos reported higher mood ratings, appreciation, energy and motivation than those asked to take photos of specific, neutral objects [80]. In another experiment, participants were split into three groups and asked to take a daily photo for three weeks of either a smiling selfie, a photo of something that makes them happy or a photo of something that would make someone else happy. When comparing the results with a control period with no photo taking activity, they found a significant increase in valence for all conditions [20, 21]. Photo taking and reminiscence was studied by Isaacs et al. [65] who displayed an increase in subjective happiness, well-being, satisfaction with life and attentiveness in those who reflected upon previously taken photos. Researchers have also studied the impact that photography has on tourists' happiness and found positive correlations between the frequency of taking photos on holiday and positive emotions and life satisfaction *if* they deemed it to be an important and mindful task rather than a stereotypical duty [51].

2.4.3 Intervention Efficacy

Lyubomirsky and Layous' positive-activity model discusses the different components of an intervention including dosage and variety [95]. It is important that the frequency of a positive activity is not "watered down" and spread out whilst also not overdone. The model also incorporates features of the participant which can influence the efficacy of the intervention. The success of the intervention can be affected by the motivation and effort that they dedicate to the task - "we argue that people need both a 'will' and a 'proper way' to gain maximal benefits from a happiness intervention" [97]. Additionally, unlike many psychological studies where the purpose of the study must be hidden from the participants, positive psychology interventions can be more effective when the participant knows and believes that the intervention will have a positive effect on them [97], which is known as intervention bias. Lastly proposing a variety of tasks within a single intervention can help to improve the overall efficacy of the intervention - "variety is the spice of self-improvement". This work by Parks et al. [125] concluded that engagement in a greater variety of tasks, in comparison to a single task, produces a greater improvement in mood.

Succinctly summarised by Eid and Larsen [39], positive psychology interventions are most effective when one dedicates time and effort to the cause, finding positivity within negativity and being mindful and open to disclosure and change. Technology-enhanced interventions have proven to be beneficial over traditional interventions in several aspects. A focus group with young adults found that they value the portability of behaviour change smartphone applications for recording and tracking behaviour and for acquiring advice and information whilst "on the go" [33]. Using a device that most people already carry with them on a daily basis is advantageous over having to carry an extra pen and paper to partake in diary-style interventions. Smartphone applications can also make use of passive sensing in order to infer specifics about a person's behaviours and use that data in order to provide context-specific intervention tasks. Even without passive sensing, smartphone notifications can act as basic reminders for participants to complete a particular action for the intervention [134]. Lastly, online social features can provide users with easy ways to share positivity, receive social feedback or get urgent access to professional help when needed [99]. Data collected from applications can also be stored online and accessed remotely by a psychiatrist or doctor, who can provide advice quickly over the internet.

2.5 Real-World Applications: Lifelogging & Affective Computing

The technologies behind tracking, monitoring and influencing emotional well-being are vital components of many real-world applications. Both active and passive data collection have importance for different applications.

The self-help movement and the field of quantified self (a.k.a. lifelogging) rely heavily on easy ways to track and monitor various aspects of one's life. Quantified-selfers are people "who diligently track many kinds of data about themselves" in order to make changes or improvements to their life upon reflection and analysis of the data [22, 105]. Widespread examples of lifelogging include fitness tracking, diet and eating habits and mood tracking; however, there are endless aspects of one's life which can be logged. The process of tracking and collecting data could be done by hand but is more commonly achieved via the use of an application. Depending on the aspect being tracked, logging could be enhanced with passive data collection using smartphone sensors, wearables or another form of automatic data collection. For common lifelogging activities, smartphone applications such as MyFitnessPal [118], for logging fitness and diet data, can provide simple user interfaces for inputting data and outputting trends, graphs and reports to help the user make decisions about what to change. There are also many applications which allow users to tailor the logging experience for their specific needs, such as the T2 Mood Tracker app [32]. For those who do not routinely track personal data, it has been shown to be more difficult to retain usage compliance without the use of monetary compensation or a prize reward [58]. This, however, was in a research scenario and those who are interested in lifelogging outside of academia are likely to be looking for a solution to a specific problem and thus compliance is unlikely to be problematic.

Similarly to the visualisation of lifelogging data to aid with decision making, visualising different aspects of mental health can have other real-world applications. In severe cases where patients are unable to verbalise or signal their emotions, either due to physical or psychological trauma or due to alexithymia, the inability to identify and describe one's own emotions [168], monitoring and sensing technologies could be extremely beneficial for health care practitioners to determine or identify the emotional state of their patients [48]. People who find it difficult to express their emotions or identify the emotions of others may also benefit from these technologies. For example, people who are affected by autism or Asperger syndrome might benefit from visual technologies which can teach or assist individuals to help them to identify emotions during communication [136]. The surfacing of emotional reactions can also be a factor for investigating usability and product testing to analyse how users or consumers feel about a particular product or design. Additionally, instant, remote access to lifelogging data can be an easy way for doctors, psychiatrists and counsellors to keep track of their patient's progress. The communication capabilities of a smartphone can also be used to provide real-time remote communication between the patient and the practitioner and the GPS can be used to keep track of the location of vulnerable patients [94].

Lastly, affective computing is the field in which computers are able to recognise and express affect in order to interact with the user on an emotional level, to better detect the user's emotions and make better decisions as a consequence [131]. Computers are able to recognise affect in a multitude of ways including facial expression, body gesture, speech and physiological signals, some of which are discussed in Section 2.3. The potential applications of affective computing are in e-learning environments to assess attentiveness, boredom, confusion, engagement and frustration which could be used to customise the learning experience based on the user's state [5, 166]. The passive detection of boredom has also been exploited in order to predict opportune times to trigger an action to engage with, such as suggesting an interesting video to watch or completing a task on a to-do list [134]. Scenarios where boredom or fatigue could be dangerous, such as when driving a vehicle, can also be an application for affective computing [180]. The ability to detect driver fatigue and issue a countermeasure can help to prevent fatal accidents. Affective video games is another field where the player's affective state can be used to modify the difficulty of the gameplay, for example. The affective state can be inferred via the gameplay itself [199] or the peripherals used to play the game, such as a gamepad [175]. Finally, social robotics, the field concerning robots which interact and communicate with humans, require technologies from affective computing to be able to appropriately detect and respond to human emotions [38]. Without these behaviours, the robots will fail to pass the Turing test, which evaluates the machine's ability to exhibit behaviours indistinguishable from that of a human. Emotion is a core aspect of human psychology and affective computing is a way to form a tighter connection between psychological state and computer systems, thus making advancements towards a seamless human computer interaction [132].

2.6 Summary

In this chapter the topic of emotional well-being was discussed, exploring the definitions of, and differences between, affect, emotion, mood, wellness and well-being. The existing literature for traditional and digital psychological assessment methods and surveys were reviewed, followed by the identification of their limitations.

Monitoring emotional well-being is important for many applications; however, traditional assessment methods are limited by participant burden, falsification, compliance, satisficing and linguistic and cultural complications. The literature review continued by describing the existing research into the passive detection of different aspects of emotional well-being which can help to alleviate some of these issues. Specifically, passive tracking using smartphone sensing, wearable sensing and online behavioural data were discussed as they are relatively unexplored areas of research. This review identified that the majority of literature using online behavioural data focus on using a single social network as a data source, and little is known about the use of data from multiple sources. Additionally, some of the literature do not take into consideration real-world ground truth data and rely on online sentiment being an accurate representation of one's emotional state. The study carried out in Chapter 3 addresses these gaps in the literature by combining data from both Facebook and Twitter and ensuring that the inferred psychological states are verified against real-world ground truth data collected via ecological momentary assessments.

This chapter then continued to review the existing research in the field of positive psychology interventions, discussing both traditional and modern interventions to aid in the improvement of various aspects of emotional well-being. The review included traditional interventions focused on life review and goal setting tasks which mostly involved a form of writing or discussion. The technology-enhanced interventions made use of smartphone features such as multimedia content, communication services and smartphone photography for their interventions. This review identified the lack of research focused on the amalgamation of traditional and technology-enhanced interventions. The study described in Chapter 4 implements the integration of momentary smartphone photography and photo review with traditional writing intervention tasks to research the relationship that such an intervention can have on mood and affect.

— 3 —

Detecting Mood Changes using Online Social Network Activity

Online social networks (OSNs) have become an integral part of our everyday lives, where we share our thoughts and feelings. Real-world communication between individuals is a multimodal experience, involving much more than the exchange of verbal information. Every day social interactions are enriched by our natural ability to interpret visual or auditory clues that help us understand the emotional state of our counterpart [4]. In recent decades a significant part of our social lives have shifted into the digital world, through the use of online social networks such as Facebook and Twitter. However, compared to real-world social interactions, OSNs do not facilitate the discovery of emotional or mood changes of the users, unless they are explicitly declared by the users themselves. Considering the significant role that OSNs play in the daily lives of individuals, we believe that there is great value in transforming OSNs into an *affective communication* medium where mood or emotional changes can be passively communicated.

In this chapter we explore the feasibility of inferring mood changes of OSN users by analysing their online activity on Facebook and Twitter. A number of previous studies have demonstrated that online activity can be a predictor for the detection of long-term psychological conditions [28] or a means for larger scale analysis of emotional trends in groups [143]. In this work we aim to identify signals from both Facebook and Twitter that can be exploited to detect individual user's mood changes within a relatively short time frame (7 days). Additionally, the majority of previous studies only analyse data from a single OSN, which we believe does not provide the best representation of real-world mood. To the best of our knowledge, many previous studies have not compared the online data with experience sampled ground truth data about their participants' psychological state. We validate our results using real-world mood data which we collect through experience sampling using smartphones.

Affective communication may involve the understanding of the psychological state of users participating in an interaction. However, considering that OSN communication is typically asynchronous and sparse, we believe that real-time emotional detection in the context of OSNs is impractical and possibly unfeasible. Instead our aim is to capture psychological mood changes that can span over longer periods of time. In this work we demonstrate that OSN activity of users can contain signals that can reveal the mood changes of individual users. We do so by exploring the relationship between online activity and actual mood changes, captured through experience sampling over a period of one month.

More specifically, we conducted a study involving 16 university students. Using online activity traces captured from their Facebook and Twitter accounts as well as self-reported daily logs of mood changes, we demonstrate that it is feasible to detect changes of mood within a window of 7 days, for 61% of the participants. We developed a machine learning classifier that can identify which users demonstrate such correlations, along with the type of correlation (positive / negative). The participants fall into three distinct groups: those whose mood correlates positively with their online activity, those who correlate negatively and those who display a weak correlation. We trained two classifiers to identify these groups using features from their online activity, which achieved precision of 95.2% and 84.4% respectively. With these results we demonstrate the feasibility of an automated technique that can discover OSN users whose activity is strongly correlated with mood change and predict the type of mood changes that can be observed, aggregated over a window of

7 days. To the best of our knowledge this is the first case of exploring links between activities from multiple OSNs and real-world mood data captured through experience sampling. This approach allows us to find evidence about correlations between online activity and real-world mood for particular types of social networking users.

3.1 Motivation

This work aims to explore the feasibility of inferring changes in the mood of individuals through the analysis of their activity on OSNs. Our aim is to develop a system that can infer mood state signals in order to enhance online social interactions. Specifically, we envisage a system that can act as a "soft sensor" which can predict mood changes through online social data. We therefore attempt to address the following research questions:

- **RQ1:** Which specific behavioural activities on Facebook and Twitter correlate with real-world mood changes?
- **RQ2:** Can activity from both Facebook and Twitter be used to infer real-world mood changes?

In order to achieve this we aim to develop a multilevel classifier that predicts the change of mood for OSN users. As seen in Figure 3.1, we consider a system that consists of: (i) a classifier that can identify OSN users whose mood changes can be predicted through online activity, (ii) a classifier that can identify how their online activity is related to their current mood, and (iii) a tracker that maps online activity to mood changes. We attempt to construct such system experimentally. Specifically, using appropriately collected datasets, we explore whether mood changes can be correlated with particular features extracted from online activities on social networking sites. The results of this analysis is then used to train the classifiers that can identify users whose mood changes can be predicted using online data.

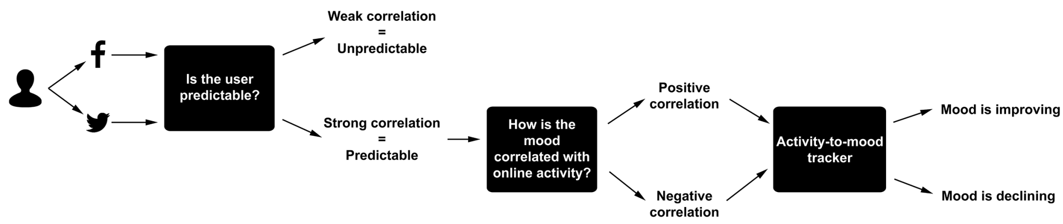

Figure 3.1: The mood tracking system. (i) User's Facebook and Twitter accounts are passively tracked. (ii) Identify whether this user's mood changes can be predicted through online activity. (iii) Identify whether their mood is positively or negatively correlated with their online activity. (iv) Online activity used to passively track mood changes.

3.2 Data Collection

3.2.1 Methodology

We conducted a study where we collected both online data and ground truth data about the mood changes of OSN users. Each participant was requested to share information with the research team about both their OSN activity and to report mood changes during the study. The study was approved by the ethics committee at the University of Kent. The literature relating to experience sampling, time- and event-based design and assessment methodology challenges in Section 2.2 was closely followed when designing this study to ensure that the best methodology practices were followed and the limitations were understood.

3.2.1.1 Consultations with Domain Experts

During the design stage, it was also important that we consulted with psychology experts to ensure that our methodology was appropriate. We held informal consultations with an academic team from the School of Psychology at the University of Kent, including a professor of social psychology and their research team. The discussion included guidance for methods of mood assessment, linguistics for the questions and assessment labels and methods to reduce participant burden. Specifically we discussed the scale, divisions, questions and labels for assessing mood, the length of the study and the time and frequency of the momentary assessments in order to keep burden and fatigue to a minimum while maintaining a good quality dataset. The following methodology was the result of these consultations.

3.2.1.2 Recruitment

The study was aimed at OSN users who maintain a relatively frequent interaction with Facebook and Twitter. Considering the general statistics about daily use of OSNs [130] we targeted younger adults between 18-25 years old who are considered more active online users. Specifically we targeted students at the University of Kent. The study was advertised through a range of University forums and each interested student was asked to submit a short survey about how often they post on Facebook and Twitter. Of the 73 people who initially registered their interest, 36 were chosen to participate in the study and 16 remained after post-study data cleaning. Those participants declared that they use Facebook and Twitter approximately once a week or more, with several answering "a few times a week", "around once a day" or "several times a day".

We ran the study during the end of the academic year (spring 2015) and over the subsequent summer break. Our intention was to ensure that students who participated would be part of the study during their exam period as well as the summer break. With this approach we anticipated that participants would demonstrate a wider variability of mood changes, possibly due to exam pressure and the relaxed summer break period that followed. Each participant was expected to be part of the study for approximately one month including part of their exam period and at least two weeks after the end of their exams. The average duration of participation was 28 days.

3.2.1.3 Ground Truth Data

The ground truth data collection required participants to install the smartphone application Easy M for Android [82] or PACO for iOS [53]. Both applications prompted the participants at 10pm every day to answer two questions: "How was your mood today, for the whole day in general?" and "How do you currently feel right now?". Participants could respond to Q1 on a 7-point Likert scale with clearly marked extremes: -3 (bad) - +3 (good) (Figure 3.2 (left)). The one dimensional input was selected to match with the commonly used Positive and Negative Affect Schedule (PANAS) to detect mood states [196]. Q2 was answered using an affect grid, initially proposed by Russell [152], in which the participant can easily record their affect on a two dimensional grid: valence (x) and arousal (y) (Figure 3.2 (right)). Through the guidance of psychology academics, the two questions were carefully worded to ensure the participants understood and could differentiate between the two answers that were required. By including questions about both their general mood throughout the day, and their affective state at the time of question, our intention was to ensure that participants would not erroneously report their current affective state as their daily mood.

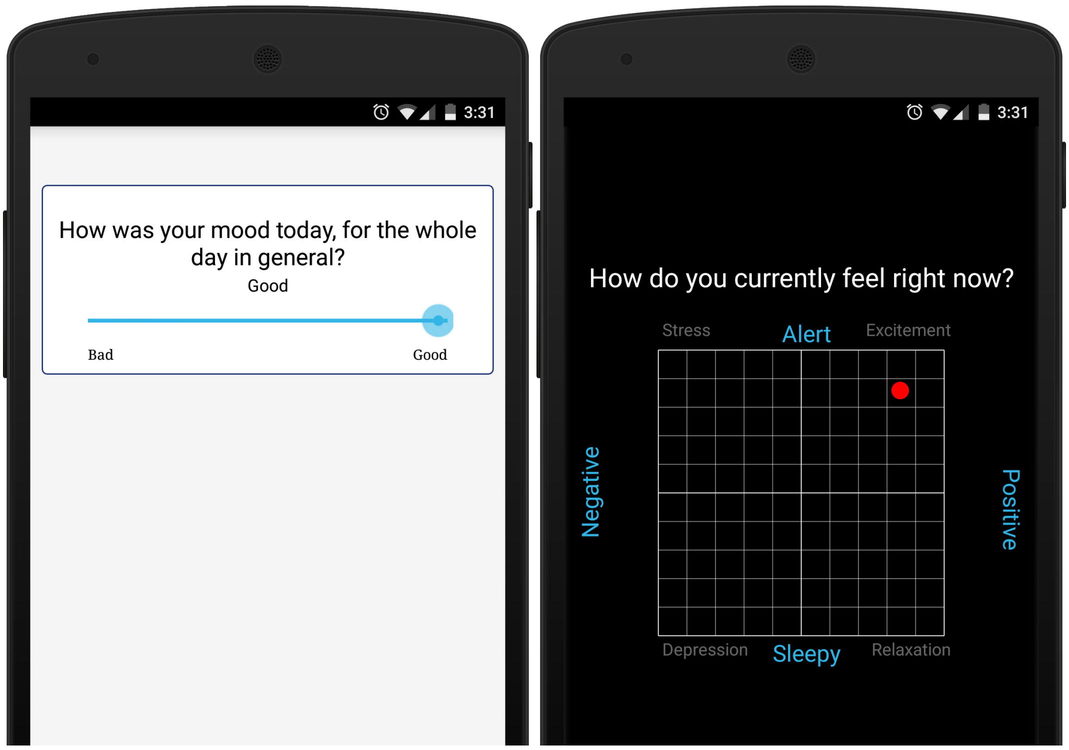

Figure 3.2: Surveying the participant's mood (left) and affect (right) within the Easy M application.

In order to maximise the response rate for the ground truth collection we tried to keep the required interaction to the absolute minimum (one notification per day). 10pm was chosen for the notification time which is before most participants are likely to be asleep. At the end of the study, the overall response rate for the ground truth was 88%.

Self-reported mood changes are considered subjective. Particular people may tend to be more positive or negative in general about their mood. To accommodate for such bias we followed the same methodology as in [76]; we standardised the reported mood using the z-score [174] as an objective metric of mood change:

$$z_{it} = \frac{x_{it} - \mu_i}{\sigma_i} \tag{3.1}$$

where x_{it} is the mood reported by the participant i on date t, μ_i and σ_i are the average and standard deviation of mood for that participant over the duration of the study.

3.2.1.4 Online Data

The online data included information about the participants' activities on Facebook and Twitter. First we set up Facebook and Twitter applications through their developer platforms. Participants were then guided to our website which prompted them to sign into our Facebook and Twitter applications with their personal accounts and grant permission to access their data. Two PHP scripts were called upon by a cron job every hour which would download activity data from each social network. For Facebook, we used their Graph API to collect data about the participant's activity from their personal timeline and home feed. These included their own posts along with likes and comments on other people's posts. It is important to note that we were able to collect personal posts regardless of the privacy settings (i.e. public, friends only) due to the API permissions provided by the participants. The script also collected select profile information such as the participant's demographics. From Twitter, the script collected all of the participant's tweets, including replies and retweets, together with their friends and followers using their API. The scripts were active for the duration of the study, collecting live data every hour which was later compiled into a single data feed for the whole duration of the study.

3.2.1.5 Data Cleaning

Following the end of the study, both datasets were cleaned. Specifically, days at the beginning and end of the study were trimmed where there were no ground truth reports submitted by the participant. Some participants unnecessarily reported their mood more than once per day, in which case the later time was used. Moreover, a number of participants were removed from the analysis using the following heuristics:

- The same mood was reported for every day of the study (straightlining).
- After trimming, the dataset was less than 15 days long.

After cleaning, the dataset consisted of a total of 16 participants including 406 individual days of data (M = 25 days per participant) and 1,760 online actions (posts, likes, etc.) performed by the participants.

3.3 Mood Detection

It is reasonable to expect that mood detection may not be possible for every OSN user. For some individuals their online activity can be more revealing of their psychological state than others. In our analysis, we tested different features, looking at the percentage of users for whom there was a statistically significant correlation ($p < 0.05$) with mood change (mood z-score).

3.3.1 Features

Using the OSN activity dataset, features were generated for each individual participant for the duration of the study. Values were calculated over a sliding window of n days, with $n-1$ days overlap. The optimal size of the window was estimated experimentally as shown in Section 3.3.2.

3.3.1.1 Sentiment Analysis

Sentiment analysis has been applied widely to discover the emotional context of messages exchanged online, including OSNs [76, 143, 191]. Sentiment analysis is the evaluation of textual data to extract useful emotional meaning. We employed sentiment analysis on the statuses and tweets that participants posted on Facebook and Twitter by using the LIWC software to extract the total number of positive and negative words from each post. For example, the phrase "Lasagne makes me happy" would return 1 positive word, 0 negative words and a total of 4 words. Through the LIWC toolkit we calculated the sentiment score S of each post using a method similar to [76]:

$$S = \frac{(n_{pos} - n_{neg})}{N} \tag{3.2}$$

where n_{pos} is the number of positive words, n_{neg} is the number of negative words and N is the total number of words in a post ($S \in [-1, 1]$).

Our null hypothesis was that the sentiment scores of the participant's posts and tweets (IV) do not correlate with their mood (DV) during the same period. For each participant we estimated the average sentiment score per post within a window of 7 days and calculated the Pearson correlation with the mood z-score over the same window. The number of participants that demonstrated statistically significant correlations ($p < 0.05$) between their mood and the sentiment score was relatively low. Specifically for the Facebook posts, 38% of the participants show correlations between mood and sentiment score, and for Twitter only 6%. An interesting observation here is the clear difference in participant behaviour on the two OSNs. The same participants appear to be more revealing of their psychological state on

Facebook than on Twitter. To some extent this is expected considering that Twitter is a public medium compared with Facebook which could be argued as being a more personal experience including friends and family. One possible explanation for the low performance of the sentiment score is the relatively sparse dataset that can be used for sentiment analysis. Within a window of 7 days, less than 1.5 days on average contain users' posts that can be analysed (Table 3.1). Twitter usage, for our participants in particular, was especially limited. Overall, we do not reject our null hypothesis and declare that sentiment score appeared to be a poor metric to identify real-world mood changes within a relatively short time frame thus we instead shifted our focus to behavioural features.

FACEBOOK – POSTS	
Posts per window (avg.)	1.4
Days without posts per window (avg.)	5.69 days
TWITTER – TWEETS	
Tweets per window (avg.)	11.2
Days without tweets per window (avg.)	6.39 days

Table 3.1: Sparsity of posts and tweets within a 7-day window. On average, almost six of the seven days are without posts and tweets from the participants.

3.3.1.2 Statistical Features

Using the datasets from Facebook and Twitter we calculated the most significant actions that a user can perform on the specific OSNs. These included counts of the following actions: Facebook status updates, likes, comments, posted links / photos / videos, Twitter posts, retweets, hashtag counts and mention counts. We consider these values as a reflection of the overall behaviour of each user on these OSNs. One important point is that the complete set of these features is significantly more dense than the raw textual posts / tweets that were considered for the sentiment analysis. This enabled us to work with denser datasets aggregated over a similar time window.

We expanded the feature set with a number of calculated / aggregate features that were derived from the raw features. Specifically, we calculated the total activity on

FEATURE	DESCRIPTION
f/tSentiment	Average sentiment score for posts and tweets
f/tAverageStringLength	Average length of post on Facebook / tweet
fActivity	No. of posts, likes, comments, posts to friends
tActivity	No. of tweets
activeActivity	No. of posts, comments, tweets and replies
passiveActivity	No. of likes and retweets
fLikes	No. of likes on Facebook
fComments	No. of comments on Facebook
fPostCount	No. of posts on Facebook
fPostsToFriends	No. of posts on friend's timelines
fTypeStatus/Link/ Photos/Video	No. of posts per type on Facebook
tHashtagCount	No. of hashtags in tweets
tMentionCount	No. of mentions in tweets
tMentionCountUnique	No. of unique mentions in tweets
tPostCount	No. of tweets
tTypeStatus/Reply/Retweet	No. of tweet per type
totalOnlineActivity	fActivity + tActivity
f/tTimeA_B	No. of posts on Facebook / tweets within a time period (e.g. fTime6_11: No. posts at 6:00–11:00)
f/tDaysSinceLastActivity	No. days since the participant was last active on Facebook and Twitter

Table 3.2: Description of features

each OSN ($fActivity$ and $tActivity$ for Facebook and Twitter), as a sum of all the counts of online actions that the user performs (e.g. posts, likes, comments, etc. on Facebook). We also calculated their overall online activity $totalOnlineActivity = fActivity + tActivity$. This is a reflection of how active each participant was online irrespective of the type of activity they performed. This was a very dense metric with values for almost every day of the study for every participant.

Motivated by the results from the work by De Choudhury et al [28] we enriched our set with features that captured the time of day that participants were active. Specifically, features $fTime0\_5$, $fTime6\_11$, etc. contain the average count of actions performed within the corresponding time windows: 00:00–05:00, 06:00–

11:00 and so forth.

Finally, we wanted to capture the level of engagement of each participant with each OSN. Based on common experience with Facebook and Twitter, we classified the different types of online actions as “active” or “passive” [34]. We classified active actions as submitting an original post or commenting on somebody else’s post, while passive actions were likes or retweets. Generally, any action that required the user to type original text was deemed as “active” and all other actions were classified as “passive”. The features $activeActivity$ and $passiveActivity$ contained the number of activities of these two classes. The aim behind these features was to explore whether more or less engaged interaction online is correlated with the mood of the user. The full featureset can be found in Table 3.2.

3.3.2 Results

We calculated Pearson’s correlation for each participant between their mood changes (DV) and each feature in our set (IV). We calculated the number of participants that demonstrated statistically significant results ($p < 0.05$) for each feature. Figure 3.3 shows the results for all these experiments. We can see that the $totalOnlineActivity$ is the feature where 61% of the participants demonstrate statistically significant correlation. These particular results were calculated over a sliding window of 7 days with 6 days overlap.

The optimal time window for the calculation of the features was estimated empirically. We ran multiple correlation tests using the percentage of users with statistically significant correlations between each feature and their mood as a metric. Increasing the size of the time window improves the performance of the $totalOnlineActivity$ feature with a plateau at the 7-day window and peak at the 15-day window (Figure 3.4). The number of statistically significant correlations is likely to increase with the window size due to the increased amount of data being included in each window; however, as our aim is to predict mood changes as accurately as

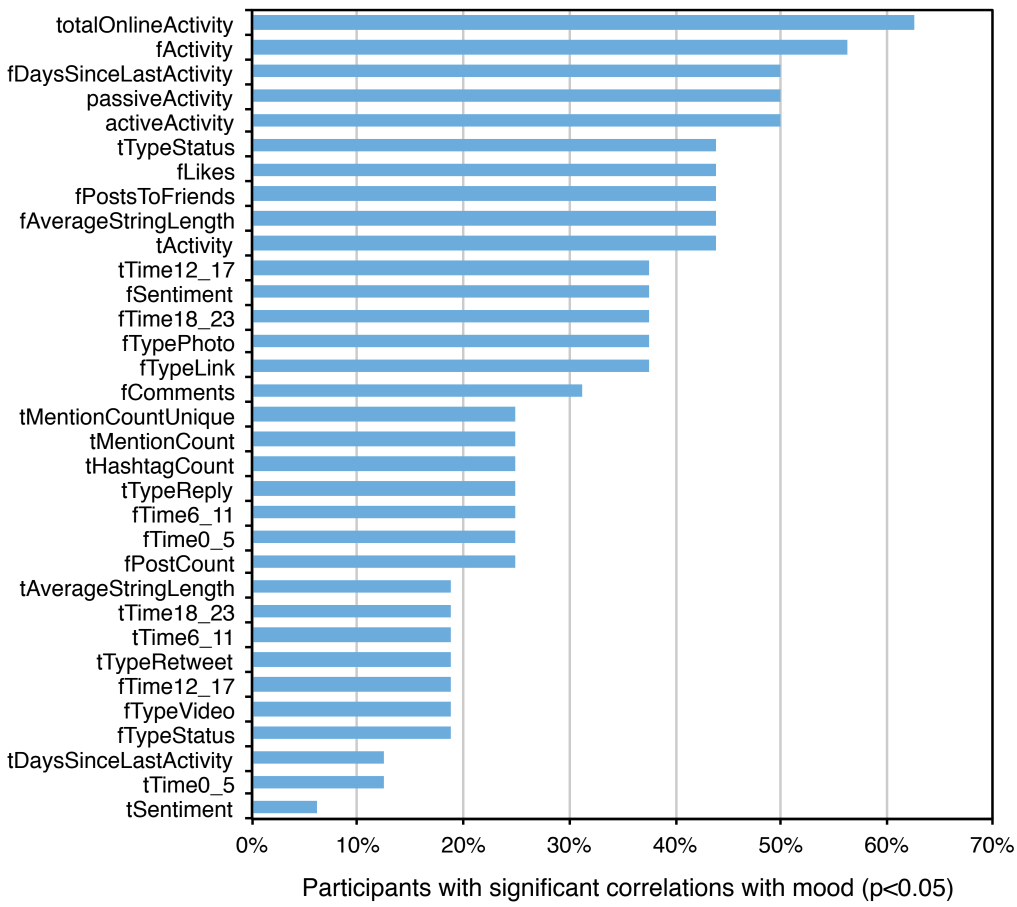

Figure 3.3: The percentage of participants who show significant correlations ($p < 0.05$) for each feature.

possible, using a 15-day window would not provide us with the desired outcome. We instead find a balance between window size and correlation performance by selecting a 7-day window for the analysis.

Based on our correlation results, we consider the *totalOnlineActivity* aggregate feature as a good indicator of mood changes online. Participants in our study showed positive or negative correlation between their daily mood changes and their overall activity online (Figure 3.5), demonstrating that when some participants were experiencing a negative mood they were more active online, while for others a positive mood was related to high online activity instead. Through informal interaction with some of our participants we received anecdotal confirmation that these results indeed match with their own perception of their online habits. Hypothetically, the causation of this phenomenon could be that people turn to online social networks

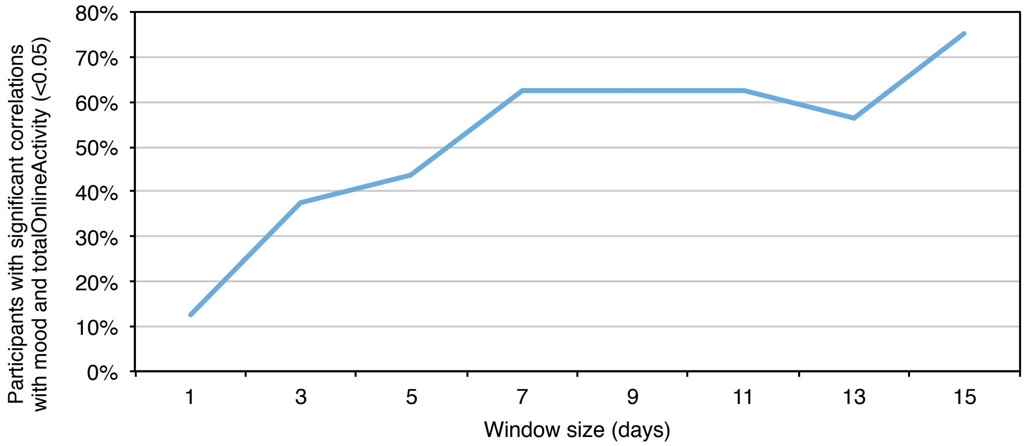

Figure 3.4: Participants with significant correlations ($p < 0.05$) between mood and totalOnlineActivity for different window sizes.

because of their mood. For example, because they are experiencing a negative mood, one may turn to OSNs to look up positive content or as an outlet to help improve their mood. In contrast, the effect could also be caused *by* the OSN itself. For example, one's mood may be influenced from interacting with the OSN, regardless of their initial intention. The specific causality is not studied in this research but could be a topic for future work.

As seen in Figure 3.6, the correlation coefficients of different users can vary significantly. However, in order to develop a practical technique to detect mood changes, it is enough to identify the signage of the coefficient of each user, relying on online activity features. Indeed, knowing that a user has a negative correlation between $totalOnlineActivity$ and mood for example, can allow the tracking of mood change (whether it increases or decreases) using OSN activity data alone.

3.4 Discovering Correlation Types

In order to automate the tracking of mood changes through OSN activity, we need a mechanism to firstly discover which users have strong or weak correlation between online activity and mood; and secondly, for those with strong correlation, discover the signage of the coefficient between their online activity and mood changes. For

Figure 3.5: Positive (top) ($r = 0.45$, $p = 0.03$), negative (middle) ($r = -0.46$, $p = 0.01$) and weak (bottom) ($r = 0.09$, $p = 0.60$) correlations between mood and totalOnlineActivity.

this purpose we developed two machine learning classifiers to detect the different types of OSN users. We relied on the feature set that is shown in Table 3.2 and we used the correlation results to define the relevant classes: *strong* vs *weak* with respect to statistical significance, and *positive* vs *negative* with respect to coefficient signage.

Firstly, we aimed to develop a binary classifier that will identify users who may

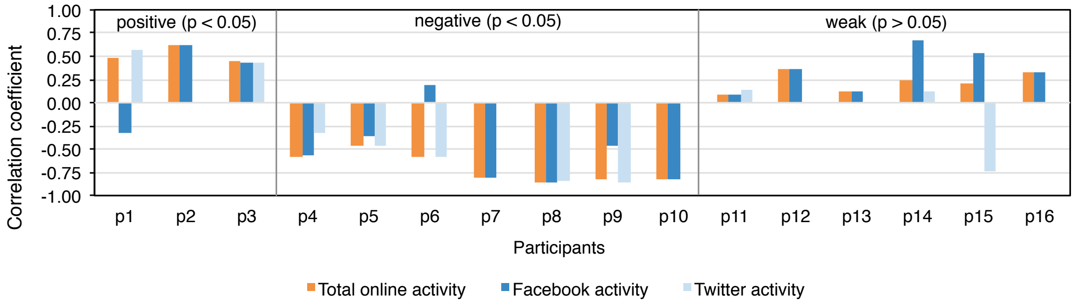

Figure 3.6: List of 16 participants with the correlation coefficients between mood and totalOnlineActivity, fActivity and tActivity. Participants on the left show a positive correlation, in the middle negative correlation, and on the right weak correlation.

show statistically significant correlation against those with weak correlations between mood and online activity. We tried to select a minimum set of features that maximised the performance of the classifiers. We followed a *hill climbing* iterative approach, where we started with the full set of features shown in Table 3.2 and then eliminated features progressively. Each step involved eliminating one feature, running the 10-fold cross validation and measuring the impact of each change to check for a positive improvement. This process is repeated for each feature until no further improvements are found. This eventually established a minimal feature set with the best performance that consists of the following features:

- $lengthFAvg$: Average length of the Facebook posts
- $lengthTAvg$: Average length of the Twitter posts
- $activePassiveRatio$: Ratio of "active" actions (e.g. new posts) over "passive" actions (e.g. likes)
- $twitterFacebookRatio$: The ratio of Twitter actions over the Facebook actions of the user.

Conceptually this feature set captures the level of commitment of different users when interacting with OSNs. Users who actively interact with OSNs by posting are inherently more engaged than those who simply observe, like or repost content.

The length of their posts and the ratio of active vs passive actions were positively correlated with how predictable their mood is according to their online activity.

Using this minimal feature set we trained two classifiers. A "Strong vs Weak" classifier allows us to discover whose online activity reflects their mood changes, and a "Positive vs Negative" classifier which identifies the type of correlation. As shown in Table 3.3 the "Strong vs Weak" classifier demonstrates 95.2% precision and 94.7% recall, while the "Positive vs Negative" classifier demonstrates 84.4% precision and 80.0% recall. We acknowledge that the design of these classifiers is based on a relatively small dataset but the high accuracy demonstrates its feasibility to be applied on a larger scale.

	STRONG vs WEAK	POS. vs NEG.
CLASSIFIER	Random Forest	Voted Perceptron
PRECISION	95.2%	84.4%
RECALL	94.7%	80.0%
F_1 SCORE	0.947	0.763

Table 3.3: Classification results. (i) Discovering users with strong correlations between their mood changes and their online activity. (ii) Classifying those with positive / negative correlations.

The combination of these two classifiers along with the use of the *totalOnlineActivity* feature allows the design of the OSN mood tracking system (Figure 3.1). Using up to one month's online data, the two classifiers can identify with high accuracy the users who can be tracked and whether they demonstrate a positive/negative correlation between their online activity and mood. When the classification is achieved, the system uses the correlation of the *totalOnlineActivity* feature across a sliding window of one week to track the mood of the user.

3.5 Discussion

The design of a system for the passive detection of mood changes can raise significant concerns for OSN users. The design of the OSN tracker was based on online data from Facebook and Twitter that the participants willingly offered and allowed to be analysed for this particular purpose. However, the main features that are used by the OSN mood tracker, can potentially be retrieved by any social contact who has sufficient access to their Facebook and Twitter streams. This implies that contacts within particular OSNs may have access to enough data to passively, and without consent, track the mood changes of their online contacts. This is a scenario that can raise significant privacy concerns and would require further exploration.

Conducting experiments with private social networks is becoming increasingly difficult, especially now after the Facebook–Cambridge Analytica scandal [9] which unveiled the malicious uses of personal data acquired by Facebook apps. Specifically, Facebook has revised their API and app review process, deprecating endpoints which they deem to be at risk and ensuring that every app is meticulously reviewed before allowing access to their users' data. The guidelines for the API also make it clear that their data must be used to provide a beneficial experience to the user, which proves to be a hindrance for researchers where collecting data is often the first step in their research. Additionally, the general public are now hyperaware of to whom they allow access to their data as a result of the news and media coverage of the Facebook–Cambridge Analytica scandal and the general consensus that people need to be more aware of their data privacy. Other online social network platforms such as Twitter still provide "open" access to their API; however, participants may still be weary of providing access to their private data. Ensuring that participants are comfortable with the research, the researchers and the privacy policy may need to be of a higher priority when conducting recruitment for data collection to ensure compliance rates are satisfactory.

The academic literature regarding methods to measure mood is not extensive, thus we consulted psychologists and used similarities from related literature in order to assess mood in an appropriate manner. As discussed in Section 2.2.2.3, mood is a single scale ranging from positive to negative and thus a Likert scale was chosen for assessment. After observing the mood tracking apps in Section 2.2.3.3, it was found that the Likert scales commonly range from 3 to 10 points. Providing too many divisions may cause too much cognitive load on the participants and lead to satisficing and too few divisions does not provide the participant with enough choice. A 7-point scale from -3 to +3 was chosen to provide a good balance of choice. The neutral zero point is important as it communicates to the participant that this is a bipolar scale. The labels at either end were chosen as "good" and "bad" to reduce the literacy requirement and to be more consistent with the casual usage of the term (e.g. "I'm in a good mood"). As discussed in Section 2.2.3, many smartphone applications assess mood using an emoji Likert scale. A purist could argue that the emojis used in those apps represent emotions rather than mood, and thus using an emoji Likert scale for mood is improper. However, a compromise must be made between forcing the correct understanding of affective phenomena and allowing users to quickly understand the meaning of the scale. I would, however, suggest the use of the basic emoji style as seen in the Feel Better app (Figure 3.7) rather than the more playful emojis used in Emotion Gram (Figure 3.8) in order to avoid ambiguity or confusion. The layout of the emojis in the Feel Better app also demonstrates the linearity of the scale, unlike in Emotion Gram.

Future work in this field might aim to use our mood tracking system in order to develop a passive lifelogging system which can automatically infer and track one's mood using their online data. This information could be kept private for personal improvement purposes or shared with a wider audience. Sharing affective state with close friends and family or with medical professionals can be especially useful for vulnerable individuals with mental health issues as an early warning system or for psychiatrists handling counselling sessions to be able to easily review a patient's state over a specific duration. Further work in the social science field might explore how sharing affective state as online content may influence behaviour towards the sharer, progressing the literature regarding online empathy. In order to realise these applications, the limitation of the sliding window size must first be addressed to ensure that mood can be tracked with more temporal precision.

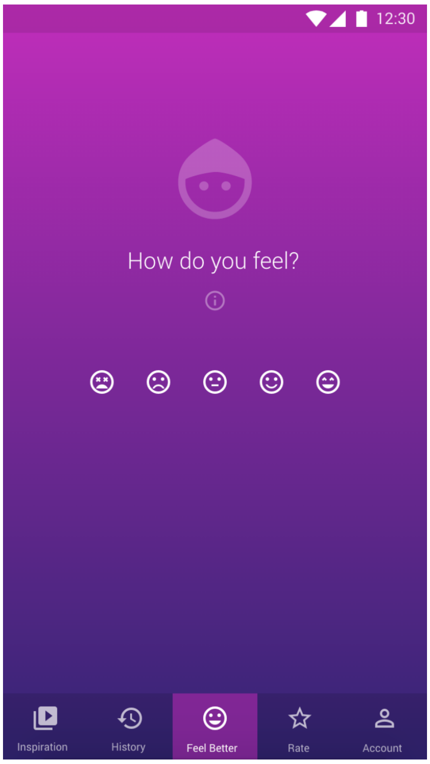

Figure 3.7: The Feel Better app uses a traditional emoji style to assess the user [112].

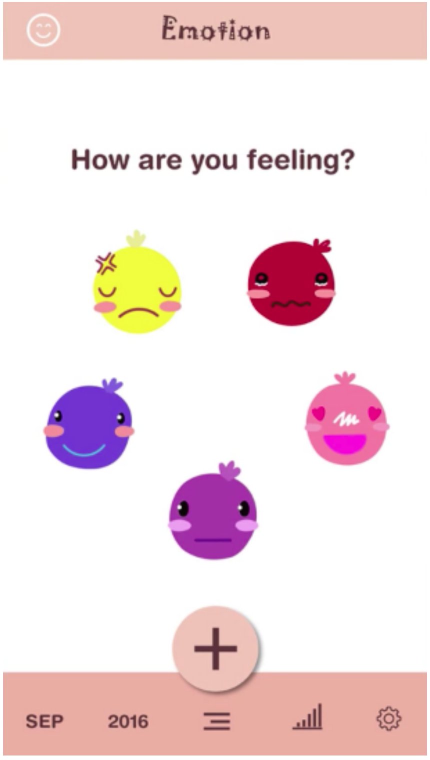

Figure 3.8: The Emotion Gram app uses more playful, cartoon-style emojis to assess the user [201].

However, the sharing of personal affective data raises many privacy and ethical concerns. A review by Sanches et al. [157] analysed the ethical approaches of existing human-computer interaction and affective health literature and presented guidance for future research directions. Specifically regarding the sharing of affective data, the authors discussed the possibility of negative behaviour towards the sharer, including the risk of discrimination due to the surfacing of a mental condition or affective state and sustaining the stigma associated with that problem. In order to maintain the positive outcomes of sharing affective data for the purpose of receiving social support, some good practices involved protecting the anonymity of the sharer's identify through de-identification or paraphrasing. For those who are particularly at risk of mental ill-health, the automated feedback provided by tracking systems may have a detrimental effect, especially when the system is likely to report frequent negative affective states. Recommended research directions involved reframing negative data in order to retain the motivation to improve their well-being and the emphasis on positive thinking and reflection upon the affective data. Lastly, relating more towards mental health conditions than affective phenomena, it is important to avoid the claim of diagnosis without the inclusion of medical expertise. Instead, the suggested approach is that of educating the participants to make sense of the data themselves and to proceed with the most appropriate actions.

3.6 Conclusions

In this chapter we have demonstrated the feasibility of detecting real-world mood changes of people through their online activity on social networks such as Facebook and Twitter. As discussed in Section 2.3.4, the majority of the state-of-the-art literature in this field focuses on using a single social network as their data source and rely on the output of their emotional inference algorithms as being the truth without making comparisons to real-world, experience sampled data. The challenges and implications of these issues are discussed in Section 2.3.5. Self-idealisation and self-presentation can skew the appearance of an online social network user, reducing the emotional congruence of their online and offline behaviours and feelings. The impact of population bias on determining how users use specific OSNs and how that can skew the reliability of data from a single social network is also discussed. Thus, trusting emotional data from a single social network without comparing it with real-world data may not yield the most accurate results. To the best of our knowledge this is the first case of exploring correlations between activities over multiple OSNs and real-world mood data captured through experience sampling. We conclude that using sentiment analysis on the written posts to detect mood is not viable due to the lack of daily posts by an average person. However, it is feasible to identify OSN users who portray their mood through their online behaviours and then monitor their mood changes using their overall online activity. We believe that these results show how OSNs can act as a medium to facilitate affective communication online and demonstrates practical ways that this can be realised.

— 4 —

Influencing Mood and Affect using a Positive Psychology Intervention Involving Smartphone Photography

Logging, assessing or inferring psychological state provides the information about one's emotional well-being; however, on its own it does not have any impact on one's life. Traditional psychological interventions involve a range of techniques and practical tasks that can influence the general mood and overall emotional well-being of a person. A common way of helping individuals improve their daily mood and satisfaction with life involves interventions where participants express their experiences and emotions through writing tasks [128]. Keeping a log of positive events in their lives, expressions of kindness and reflection on happy memories are common techniques that have shown positive results in improving overall mood and feelings of happiness.

Motivated by these techniques, we developed SnapAppy, a smartphone application to integrate momentary photography with traditional intervention methodologies to conduct a month-long positive psychology intervention. The study investigated the feasibility of using smartphone photography as an intervention to help improve the emotional well-being of individuals, specifically their mood, affect and satisfaction with life. We aimed to explore the effect of using smartphone photography as a method for capturing happy memories and positive events in the daily lives of participants, and enabling the review and reflection of such events through photos.

With the wide use of smartphone devices, photography has become a prominent activity for users. The popularity of photo sharing applications like Instagram and Snapchat has created a wide community of primarily younger smartphone users where smartphone photography is a common daily hobby. In this work our motivation is to leverage the familiarity of these user groups with smartphone photography, in order to develop an easy-to-use intervention that can have a positive effect on their mood, affect and satisfaction with life. Specifically, we developed SnapAppy to motivate users to take photos of positive events in their daily lives. Through this smartphone app, users can annotate, review and reflect on the photos they take. We consider the task of taking and reviewing photos as a translation of the traditional writing exercises that are applied in typical psychological interventions. Considering that, with the proliferation of technology, daily writing tasks are becoming less common for typical users, our objective is to explore a more familiar method for such positive psychology interventions. Familiarity with the particular technology can lead to higher adoption and retention of participants.

We deployed the application across a group of 74 participants who used the app for an average of 35.3 days. As part of the study, participants submitted daily mood reports and regular surveys of affect and satisfaction with life. Furthermore, the app was instrumented to allow us to collect detailed activity logs of how the users interacted with the app on a daily basis. Through the analysis of the collected datasets, we identified correlations between the level of engagement that participants had with the app and the captured change in their overall psychological state. Moreover, an analysis on the contents of the photos that were captured showed that particular combinations of photo content seem to have a more positive effect on the change of the participant's mood and affect. The results indicated that features including the number of photos, the spread of categories, the effort applied to annotating photos, the number of photos revisited and photos of people were positively correlated with an improvement in the participant's mood and positive and negative affect.

4.1 Motivation

Frequent, spontaneous smartphone photography is a popular activity with most smartphone users. Photography in general is a form of memory augmentation as well as a way of expressing a person's feelings and attitudes. Drawing similarities with the practice of personal diaries, we consider that smartphone photography can play a similar role when presented as a form of intervention to help improve a person's emotional well-being. In this work we aim to translate the techniques applied in positive psychology interventions, where the participants are required to use writing in order to record and recollect positive events in their lives. As one of the first attempts to explore the practice of smartphone photography as a potential positive psychology intervention, our primary aim is to see if there is indeed an association between the practice of smartphone photography and changes in mood and affect. Our primary objective is twofold: firstly to explore if smartphone photography is correlated with positive changes in mood, affect and satisfaction with life and secondly whether there are particular aspects of photographic activities (i.e. photo taking frequency, contents of photos) that demonstrate a correlation with the effectiveness of the intervention.

Specifically, in the context of a smartphone photography positive psychology intervention, we attempt to address the following research questions:

- **RQ3:** Do smartphone photography activities correlate with changes in mood, affect and satisfaction with life?
- **RQ4:** Which specific smartphone photography activities show a statistically significant correlation with change in mood, affect and satisfaction with life?

In attempting to answer these questions we conducted a study involving a smartphone application called SnapAppy that was developed to facilitate a photo taking intervention associated with changes in daily mood and regular changes in affect and satisfaction with life. Following a study involving 74 participants for a period of a month, we explored detailed activity logs to identify possible correlations between reported psychological states and the participants' behaviour.

4.2 Methodology

A user study was conducted between summer 2017 and spring 2018 which was approved by the ethics committee at the University of Kent. A key component of this study was a smartphone application that was purposely designed to allow us to explore the effects of smartphone photography as a form of positive psychology intervention. The literature regarding traditional and digital psychological assessment, experience sampling (Section 2.2), existing positive psychology interventions and intervention efficacy (Section 2.4) were closely followed when designing this study to ensure that the best methodology and intervention practices were followed and the limitations were understood.

4.2.1 Consultations with Domain Experts

During the design stage, it was also important that we consulted with psychology experts to ensure that our methodology was appropriate. Similarly to the study in Chapter 3, we held informal consultations with an academic team from the School of Psychology at the University of Kent, including a reader in psychology with expertise in emotion, addiction and attention and their research team. The discussion focused on presenting the proposed app and methodology to confirm that it was a viable intervention and used appropriate methods within the field of psychology. Specifically the existing traditional intervention literature, linguistics for the intervention tasks, the chosen psychological surveys and methods for participant retention were discussed. The following sections were the result of these consultations.

In addition to the consultations with psychology experts, we also conducted a small, informal user trail with nine people within our faculty in order to identify functionality, design and user experience issues and programming bugs. This is a common practice during app development and is often conducted multiple times during the design and development process. We conducted this trial on a prototype app once the majority of the features were functional. The users were instructed to download and use the app for a few days before being asked to convene to discuss their feedback. We also passively observed some people using the app in order to identify usability problems. Following the feedback, the identified functionality, design and usability issues and programming bugs were fixed before running the intervention. Co-participatory design methods were not used because the main objective was to design the app to meet the guidelines from the domain experts and for it to be used as an instrument to address the research questions.

4.2.2 The Application

A smartphone application called SnapAppy was developed for Android [84] and iOS [85] smartphones using Apache Cordova [3]. Using Cordova, SnapAppy was developed as a hybrid app using web technologies and compiled for both platforms, saving time and resources. The main functionality of the app was to support photo taking and annotation and the reviewing of those photos. To allow the app to function as a form of positive psychology intervention, additional features were added allowing participants to categorise photos into different types of positive moments or actions that the photos may represent. In general, the app offers the following main functions:

4.2.2.1 Photo Taking and Annotation

Participants can use the app to take photos using the smartphone camera or upload a photo from the device. For each photo the participant can assign a category that corresponds to the type of positive situation the photo represents. These categories help to guide the participant, providing inspiration for adding a variety of photos into the app. The category definitions are displayed when selecting a category to

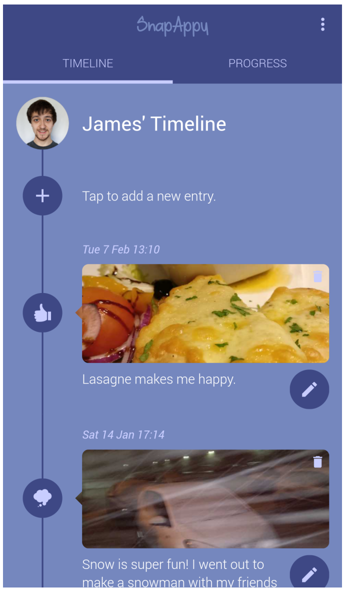

Figure 4.1: Entry photos, categories and descriptions can be viewed on the timeline screen.

Figure 4.2: When taking a photo for a new entry, it must be assigned to a category corresponding to its type of positive experience.

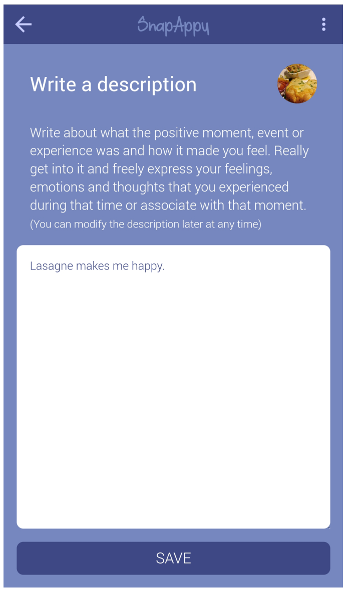

Figure 4.3: After assigning the photo to a category, one can write a description describing how the event made them feel.

ensure each participant has a consistent understanding, as shown in Figure 4.2. The possible categories that participants can assign to their photos are based on the types of activities used in traditional psychological interventions such as writing about emotional experiences, reminiscing and recollecting positive memories, counting kindnesses, expressing gratitude and writing about events considered to be caused by destiny or fate (Section 2.4). They were selected from a broader list of traditional concepts which also included goal setting [165], planning skills [102], unresolved conflict [127] and pent-up emotion [72]. These concepts were excluded due to the incompatibility of integration into the photo taking and annotation design. Finally, the five categories that were provided were:

- Positivity: A positive moment, event or experience that is happening right now. [18, 128]
- Reminiscence: A positive moment, event or experience that happened in the past. [24]

- Kindness: An act of kindness, performed by yourself or someone else, or that you have received. [120]
- Gratitude: An entity that you are grateful or thankful for (e.g. family, friend, pet, job). [43, 97]
- Destiny: A positive moment, event or experience that might never have occurred if something in the past had happened differently (e.g. never meeting a particular friend if a particular job wasn't taken). [74]

For each photo, participants had the option to write a description about what the positive moment, event or experience was and how it made them feel. Participants were encouraged to freely express their feelings, emotions and thoughts that they experienced during that time or associated with that moment (Figure 4.3). No other photos were captured outside those taken in, or uploaded to, the app.

Immediately after logging in (Figure 4.4), the participants were familiarised with the study through the use of multiple onboarding screens explaining what was required of them (Figures 4.5 – 4.8). These screens were also accessible at any time during the study via the "Help" button within the menu.

4.2.2.2 Photo Timeline and Review

The main entry screen of the app is a timeline containing the photos the participant has taken (Figure 4.1). The participant was free to scroll through their timeline at any time to revisit entries that they had previously logged. Tapping on an entry's photo would maximise it to be viewed fullscreen. Photo reviewing also allowed participants to change both the category assigned to that photo and the text description.

4.2.2.3 Psychological Surveys

In addition to logging daily entries, participants were notified to log their mood at the end of each day. Due to the difference in everyone's circadian rhythm, the

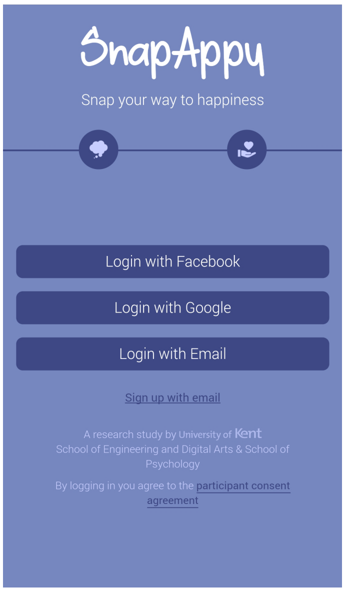

Figure 4.4: Participants can log in to the app using either Facebook, Google or an email and password.

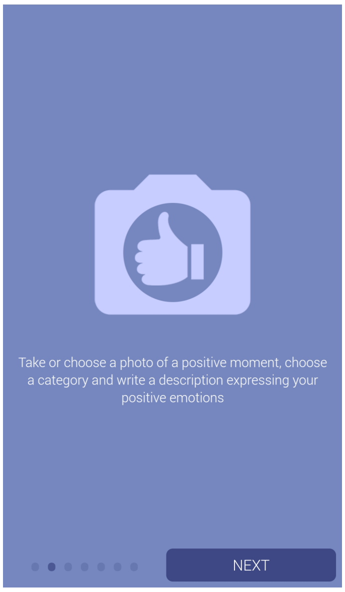

Figure 4.5: During onboarding, participants are instructed to take or choose photos of positive moments.

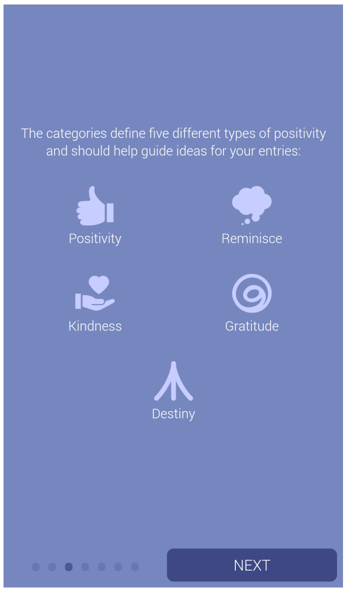

Figure 4.6: Participants are introduced to the five photo categories.

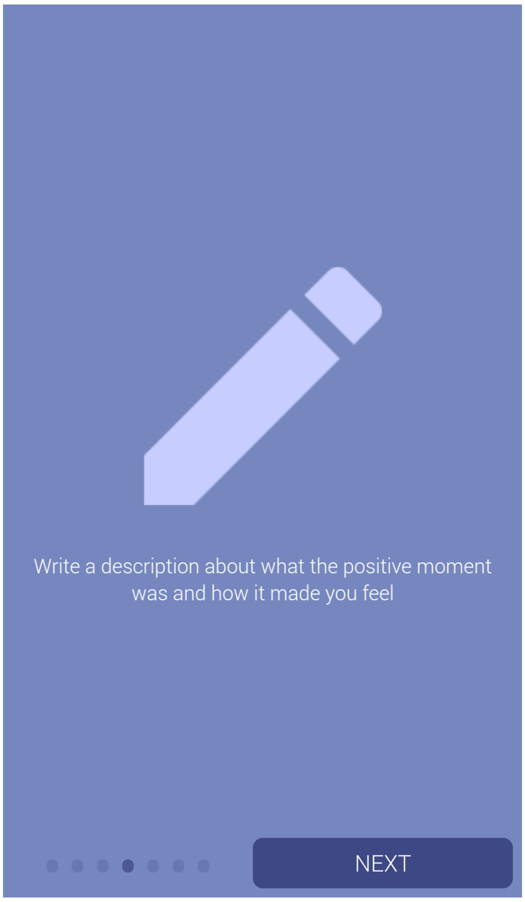

Figure 4.7: Participants are advised that they will need to annotate photos with an emotional description.

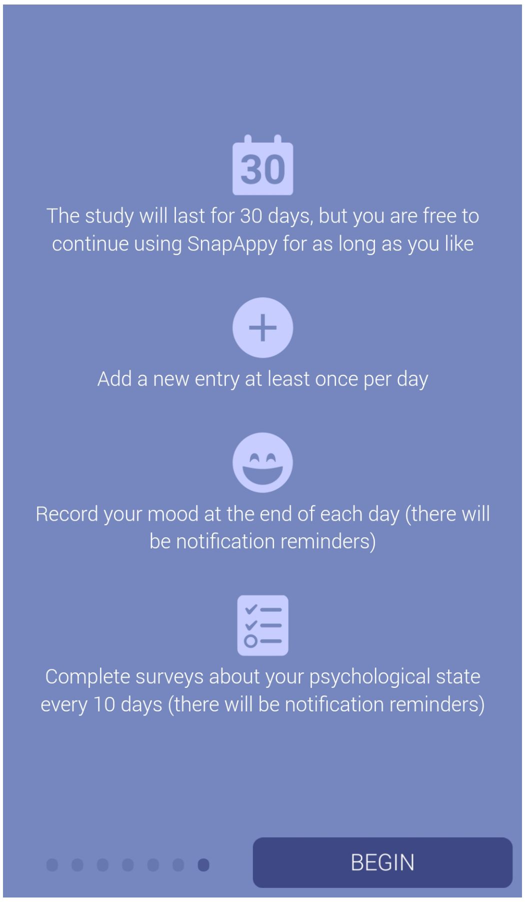

Figure 4.8: Finally, the participation instructions for the study are listed.

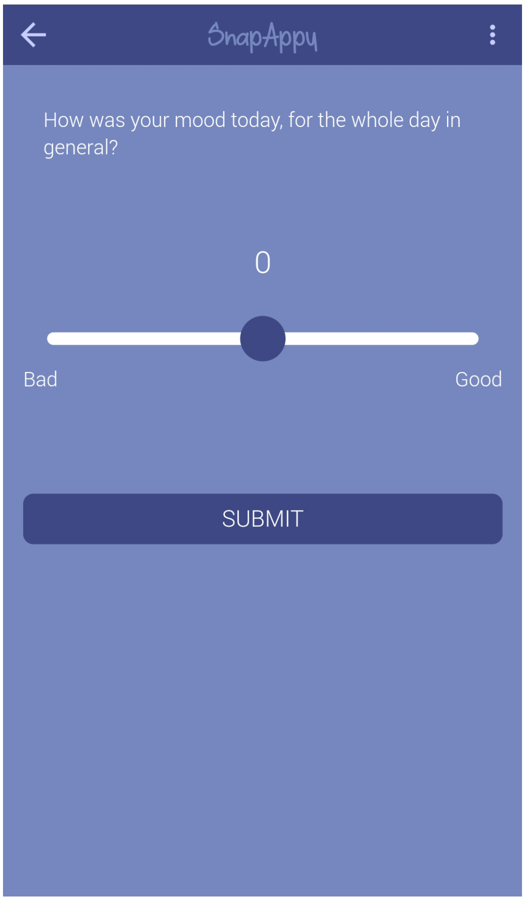

Figure 4.9: Mood is reported using a 7-point Likert slider ranging from -3 (bad) to +3 (good).

Figure 4.10: The survey screen showing one affective state and the various answers from the PANAS survey.

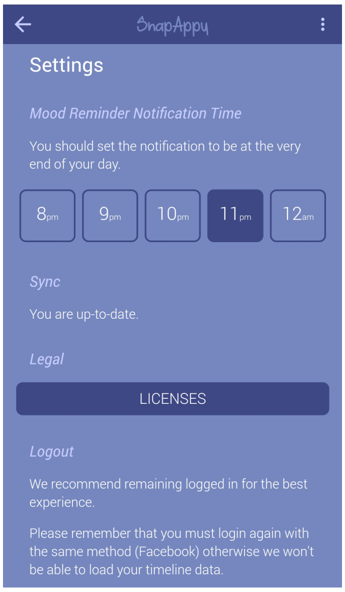

Figure 4.11: Participants could change the time of their mood notification in order to suit their sleep schedule.

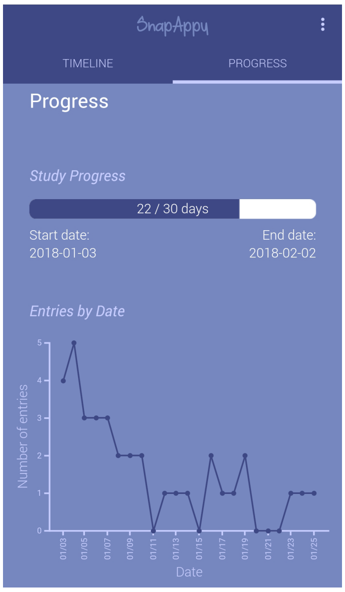

Figure 4.12: This screen shows the participant's progress through the study and the number of entries per day.

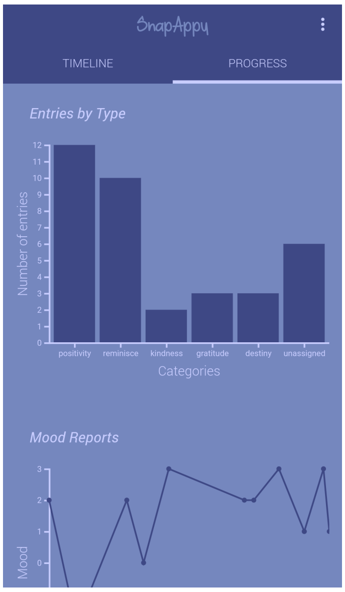

Figure 4.13: The progress screen also shows the number of entries per category and the participant's mood reports.

time of the notification was not fixed. Instead, the default was set to 10pm and we offered the choice to change it between 8pm and midnight, which allowed the participants to choose an appropriate time according to their sleep schedule (Figure 4.11). This level of freedom will not be appropriate for all study designs and may, in some cases, introduce an extra variable into the analysis. However, in our case, it ensured that emotional events occurring late in the day were incorporated into the daily mood reports. Participants were asked "How was your mood today, for the whole day in general?" and could respond on a 7-point Likert scale with clearly marked extremes: -3 (Bad) to +3 (Good) [86] (Figure 4.9).

Furthermore, in order to capture more longer-term effects of the intervention, the app invited participants to submit formal surveys approximately every 10 days. We incorporated commonly used measures to assess the psychological state of participants before, during and after the intervention. Specifically, we chose the Positive and Negative Affect Schedule (PANAS) [196] and the Satisfaction with Life Scale (SWLS) [36] to be most appropriate for this study. Both metrics offer objective, normalised values about the psychological state of each participant and are favoured by traditional psychological studies instead of more subjective assessment methods such as self-efficacy and the perceived effectiveness of an intervention. These two metrics are the most commonly used measures in both Sin and Lyubomirsky's meta-analysis of positive psychology interventions [169] and our own literature review. The PANAS comprises of two affect scales, one that measures positive affect and the other that measures negative affect. The positive and negative scales consist of ten emotionally descriptive words such as "excited", "scared" and "inspired". The extent to which the participant feels each word must be responded to using a 5-point Likert scale ranging from "extremely" to "very slightly/not at all" resulting in scores ranging between 10–50. For the positive scale a higher value indicates a more positive affect; for the negative scale a higher value indicates a more negative affect. The SWLS metric consists of five statements in which the participant must respond to on a 7-point Likert scale ranging from "strongly agree" to "strongly disagree". The answers produce a single value between 5–35 where a higher value demonstrates a higher satisfaction with life. PANAS and SWLS were also chosen for their relatively quick completion times as the in-the-wild nature of this study would not have been conducive for longer surveys. Furthermore, using shorter surveys meant that we could conduct psychological assessments *during* the study, rather than merely pre-test and post-test as was common in past literature. Therefore every 10 days during the study, participants were notified to complete PANAS and SWLS surveys (Figure 4.10). From herein we will refer to these collectively as "surveys". On day 0 they also provided information about any other strategies or practices for influencing mental health in which they were participating. None of the participants reported that they were engaging in any strategies which might interfere with the study, thus no exclusions were deemed necessary. The participants were not screened for pre-existing mental health conditions and no major real-world events were recorded that may have caused an emotional effect on the whole population of participants.

4.2.2.4 Retention and User Engagement

Although not strictly part of the main intervention, the app was instrumented with a few screens that allowed participants to observe their level of participation in the study. A "Progress" screen in the app allowed users to track their start and end dates for the study and view graphs showing information about their entries, categories and mood reports (Figure 4.12). The graphs were designed as a form of feedback for the participants in order to increase engagement; in particular the bar graph comparing the number of entries belonging to each category was designed to encourage the participants to experiment with the different entry categories [125] (Figure 4.13).

Finally, the app was instrumented with a detailed logging system that captured all interactions within the app. This included every "tap" and "scroll" gesture the participants performed, the time they spent performing specific actions, the amount of text they entered, all the photos they took and how they interacted with them.

4.2.3 Recruitment & Exclusion Criteria

Recruitment was aimed at smartphone users who could read and write in English. The study was advertised publicly on social media platforms including mental health related Facebook groups and Subreddits and also within the University of Kent via forums and well-being events. The study was aimed at a general population rather than a specific clinical population because, as stated in the literature, positive psychology is not focused on providing solutions to negative mental health issues. The advertisement campaign invited participants to download the smartphone app from the App/Play Store onto their personal smartphone. Enrolment in the study was performed through the smartphone app. Every participant was compensated with course credits if they were a student or entry into a prize draw of £50.

The recruitment campaign resulted in 201 people who downloaded the app and 87 people who started the study (by taking at least one photo, completing the first survey period and logging at least one mood report). Each participant was expected to engage with the study for a minimum of one month; however, they were encouraged to continue using the application if they so desired. Participants were asked to log at least one entry (photo) per day about a positive moment, event or experience. In addition to logging daily entries, participants were notified to log the two different types of psychological ground truth outlined in Section 4.2.2.3. The dropout rate for the study participants was fairly low. Figure 4.14 shows the initial 87 participants who remained engaged across the duration of the study. 33 participants continued using the application beyond the 30 days of the study and only 6 participants engaged for less than half of the required duration. Due to the "in-the-wild" distribution of the app, users will often install an app out of curiosity rather than having a genuine intention to participate in the study, which could explain some of the early drop-outs from the study.

After the completion of the study, the cohort of participants were further assessed in terms of the quality of the submitted data. In order to be able to analyse the

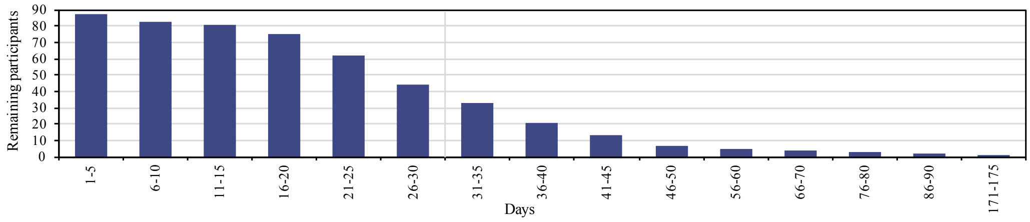

Figure 4.14: Distribution of participants over time. The recommended duration for participants to engage with the study was 30 days. A number of participants continued using the smartphone application beyond that duration.

collected datasets it was necessary for each participant to have submitted sufficient ground truth data. A set of exclusion criteria were used to remove participants who did not submit enough data. Participants or survey periods were excluded from the analysis using the following criteria:

- Participants who only completed 1 or 2 surveys.
- Participants who reported <7 daily mood reports in total (approx. $\frac{1}{4}$ of the 30 day study duration).
- Participants who reported exactly the same mood every day (straightlining).
- Periods which were <5 or >20 days between surveys (this occurred when participants missed a survey or due to an error).
- Days where no mood was reported.
- If the participant reported their mood more than once per day, the later time was used.

The resulting dataset included a total of 74 participants who were aged between 18-52 and 63 were female. The average age was 21 alluding to the fact that this form of intervention may be more popular for younger users who are more adept with the use of smartphones and spontaneous smartphone photography.

4.3 Descriptive Statistics

The average duration of participation was 35.3 days. During that time a total of 2,178 photos were taken, with an average of 29.4 photos per participant across the study (per day: min = 0, M = 0.83, max = 8). 52% of the photos were assigned to the positivity category, 16% were reminiscence, 13% gratitude, 7% kindness, 4% destiny and 8% remained unassigned (Figure 4.15). Descriptions had an average length of 93.3 characters (min = 0, Mdn = 57, max = 973) (approx. 18 words), participants spent a median of 32.7 seconds writing them and 190 photos had no descriptions. On average, participants revisited their old photos 1.1 times per day with a range of 0–32 times. Participants logged a total of 2,535 mood reports with an average of 34.3 per participant and 240 surveys with an average of 3.2 per participant. These results show significant variation in the levels of engagement across our participants. This variation enabled us to explore how different levels of engagement may demonstrate potentially different associations with mood or affect as reported by the participants.

The distribution of mood reported by all participants during the study is shown in Figure 4.15. The majority of moods reported were >0 (M = 1.18) indicating that participants reported to be feeling positive more often than negative. SWLS scores range from 5–35 and the distribution shows that the scores have an average of 22 which is within the average range (20–24) for this metric. PANAS scores are calculated separately for positive and negative affect and scores range from 10–50. The distribution of positive PANAS scores are relatively spread, with an average of 26 which is slightly lower than the mean score of 31.3 found by Crawford and Henry [25] who reported average scores for the general adult population. The negative PANAS scores have an average of 19 which is slightly higher than the mean score of 16. These scores demonstrate that the participants overall followed a similar distribution of affect that can be expected by the general population. The relatively lower than average scores for the positive PANAS could indicate a possible selection bias for our participants where people with lower positive affect might be more

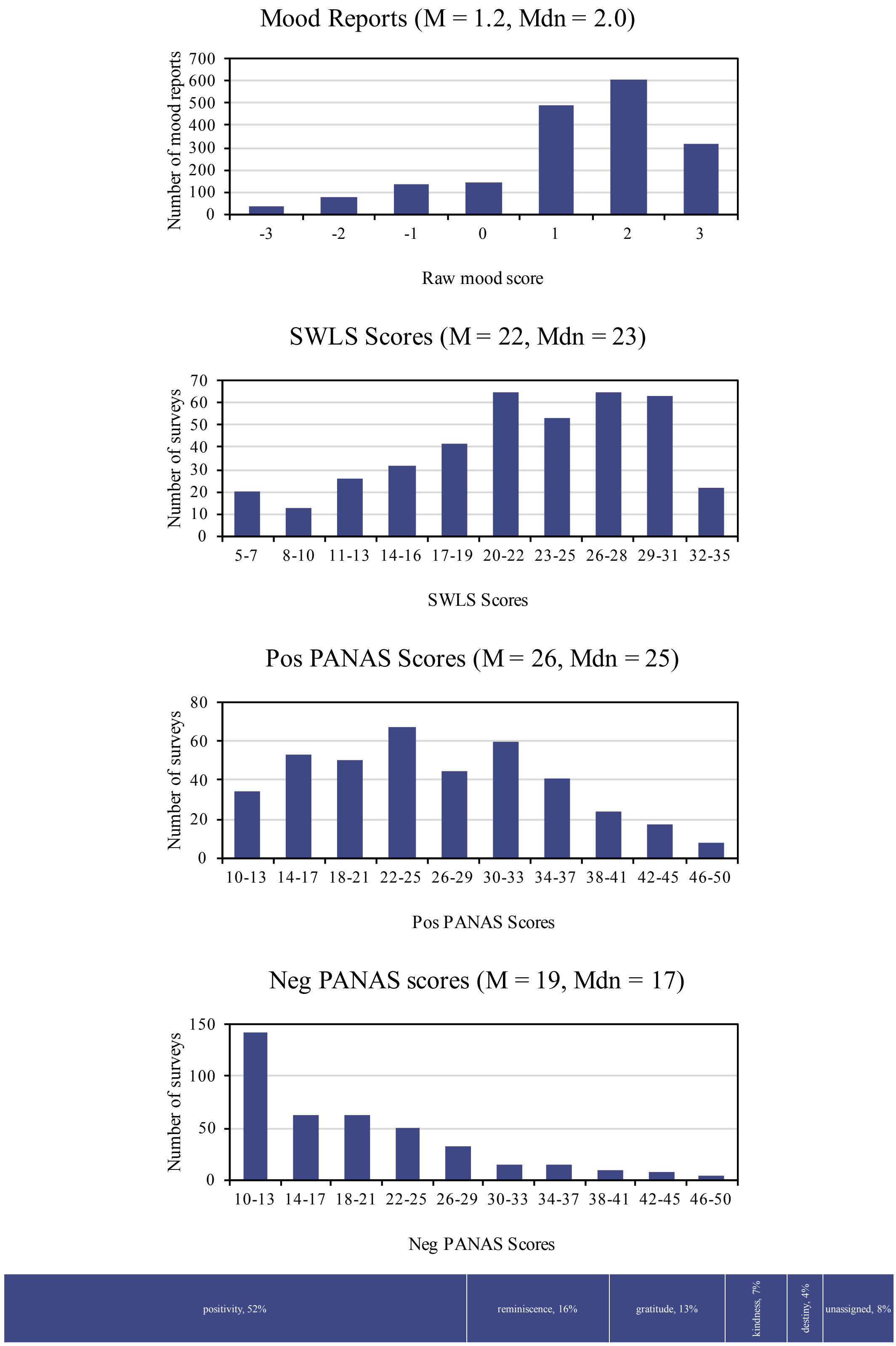

Figure 4.15: Distribution of mood reports, SWLS and PANAS scores and entry categories.

interested in joining a positive psychology intervention. However, as our objective is to discover the possible relation that a photo taking intervention may have on their affective state, we need to explore whether their level of engagement with the intervention is related to their change of psychological state during the study.

4.4 Inferential Statistics

In our analysis, we attempt to address the main research questions by analysing the datasets collected during the study. Our main approach is to explore how the participants' activities using the smartphone application may be related with the observed changes in their daily mood and affective state throughout the study.

4.4.1 Before–After Analysis

We began our analysis by looking at the overall effect of the intervention, attempting to see if there was a common pattern of change across the whole cohort. The null hypothesis was defined as no significant change in the participant's survey scores (DV) between the start and end of the study. Ignoring any details about what the participants did during the study, we tried to see if the overall affective state of the participants changed after their participation by using the formal surveys that they completed at the beginning and end of the study. To explore if there was a statistically significant change, we performed a Wilcoxon signed-rank test on the survey scores. With respect to the PANAS results, the test showed that simply participating in the study did not lead to a significant difference in the positive PANAS scores (before $Mdn = 25$, after $Mdn = 24$, $r = .032$, $p = .609$, $T = 181.5$, where T is the test statistic). Similarly, no significant change was found for the negative PANAS scores (before $Mdn = 19$, after $Mdn = 18$, $r = -.011$, $p = .855$, $T = 132$). When analysing the SWLS results, we did find a statistically significant change in the participants' satisfaction with life scores, with a 10% median increase out of the maximum range of scores (before $Mdn = 20$, after $Mdn = 23$, $r = .192$, $p = .002$, $T = 257.5$). Additionally, the box plots in Figure 4.16 show the raw scores for positive PANAS, negative

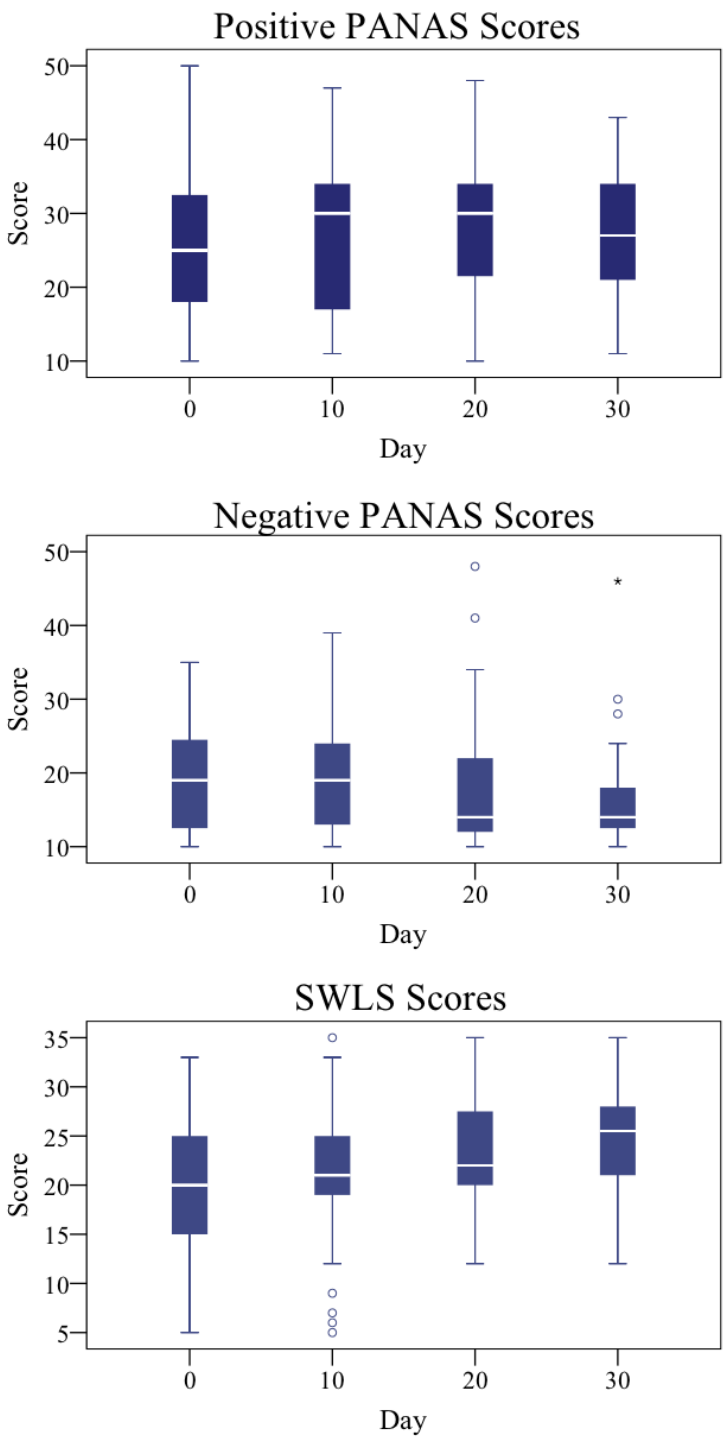

Figure 4.16: The raw positive PANAS, negative PANAS and SWLS scores across the duration of the study.

PANAS and SWLS across the duration of the study. Although the average survey results showed improvements over time, we may observe a slight plateau of the average PANAS scores between days 20 and 30. This appears to approximately align with the data in Figure 4.14 and may suggest a "sweet spot" for this type of intervention, perhaps due to participant burden, fatigue or dropouts which can often be experienced with longitudinal studies. There were also a significant range of scores across the intervention for all surveys. This demonstrates that simply joining the study would not necessarily lead to a clear improvement in the participants' positive and negative affect or satisfaction with life and we therefore do not reject the null hypothesis. However, these results gave us the opportunity to investigate differences in behaviours between our participants and identify particular patterns that could potentially lead to either a better or worse outcome.

4.4.2 Analysis per Survey Period

In our study we had two different types of ground truth. Participants were required to complete surveys every 10 days during the study, which allowed us to see significant changes within such periods. Moreover, participants submitted short daily mood reports, which gave us an indication of how mood changed on a daily basis. The regular survey reports formed a natural split over our dataset into periods where we can study changes across a number of days. Although the app notified participants every 10 days to submit a survey, participants did not always submit their surveys when they were notified, thus participants who submitted too few surveys were ignored. Following the exclusion criteria discussed in Section 4.2.3, we were left with a total of 134 periods across 53 participants (out of the total 74 participants who completed the study). Each participant generated an average of $\sim$3 periods of data (encompassed by 4 submitted surveys) and the average period duration was 10.9 days (SD = 2.7).

In our analysis we tried to explore how the engagement of the participants with the smartphone application could be related to the changes of psychological state during each period. The null hypothesis was that specific behavioural features (IV) will not correlate with the change in psychological state (DV) during that period. Using the logs generated by the application we extracted a number of features as the independent variables which represent how users engaged with the application. These included the number of photos taken per day, how often they scrolled on their timeline or viewed particular photos in full screen, and the sentiment score of the photo descriptions. Table 4.1 describes the full list of extracted features. These features are translations or interpretations of existing research discussed in Section 2.4 such as writing tasks, intervention efficacy and photo-related interventions. This therefore reduces the likelihood of a type I error and we expect the null hypothesis to be rejected for the majority of these features.

FEATURE	DESCRIPTION
Photos taken	Average number of photo entries per day.
Photo entry consistency	The spread (low standard deviation) of photo entry counts across the days during the period.
Category spread	The spread (low standard deviation) of photo entries across the five categories.
Avg. desc. length	Average character length of the photo descriptions.
Avg. desc. duration	Average number of seconds spent writing the descriptions.
Desc. edits	Average number of times the descriptions were edited per day.
Avg. desc. positivity	Average sentiment score of the entry descriptions calculated using LIWC [92]. *(No. of positive words - No. of negative words) / All affect words*
Photos viewed	Average number of times photos are tapped on to be viewed fullscreen per day.
Timeline explored	Average number of times the user scrolls up or down the timeline (where the scroll length is more than the height of three entries and with an interval of >10 seconds) per day.

Table 4.1: Description of features.

With respect to the ground truth, the psychological features were extracted by calculating the change in positive PANAS scores (ΔP_{pos}), negative PANAS scores (ΔP_{neg}) and SWLS scores (ΔS) between the survey responses encompassing the period. The average mood during that period was also calculated and defined as $\bar{M}$. Mood reports were standardised within participants using z-score so that the value represents the deviation from the participant's average mood across the whole study:

$$z_{it} = \frac{x_{it} - \mu_i}{\sigma_i} \tag{4.1}$$

where x_{it} is the mood reported by the participant i on date t, μ_i and σ_i are the average and standard deviation of mood for that participant over the duration of the study. These values form the dependent variables.

4.4.2.1 Results

We explored potential correlations using Pearson's correlation coefficient between each feature and our ground truth for the 134 periods. The results are shown in Table 4.2.

We observe that the *average description length* has positive and significant correlation with ΔP_{pos} (r_p(133) = .2063, p = .0168) showing that there is a relationship between the length of the entry descriptions and an improvement in positive affect. These correlations fall in line with past literature which describes that the more effort the

FEATURE	ΔP_{pos}	ΔP_{neg}	ΔS	$\bar{M}$
Photos taken	0.0083	-0.1128	0.0070	**0.2418****
Photo entry consistency	0.0868	**-0.1698***	0.0292	0.1619
Category spread	-0.0699	-0.0848	0.0327	**0.2353****
Avg. desc. length	**0.2063***	-0.0023	0.0480	0.1005
Avg. desc. duration	0.0669	-0.1305	0.1165	0.1199
Desc. edits	0.1089	-0.1462	0.0352	0.1481
Avg. desc. positivity	-0.0638	0.0006	-0.0368	0.0959
Photos viewed	-0.0032	**-0.1770***	0.0723	**0.2048***
Timeline explored	-0.0570	-0.1630	0.0032	0.1075

Table 4.2: Correlation results of survey features. $^{*}p <$ **.05**, $^{**}p <$ **.01**

participant puts into the intervention, the more effect it has [97].

ΔP_{neg} shows a significant, negative correlation with *photo entry consistency* (r_p(133) = -.1698, p = .0499). Past literature has shown that a consistent, distributed use of an intervention is key to its success [95] which reflects this correlation. The number of *photos viewed* is also negatively correlated with ΔP_{neg} (r_p(133) = -.1770, p = .0408) showing that revisiting old photos is related to a lower negative affect. The correlation with *timeline explored* is also close to $p < 0.05$ which is logical considering reviewing old photos requires scrolling through the timeline.

ΔS does not show any significant correlations with any of the features. This could be due to the fact that SWLS is a more long-term assessment of the participant's state, thus attempting to observe changes over shorter periods is not appropriate. This result also suggests that the significant improvements in SWLS found from the Wilcoxon signed-rank test (Section 4.4.1) may not be related to the way participants engaged with the intervention.

$\bar{M}$ shows significant, positive correlations with *photos taken* (r_p(133) = .2418, p = .0049), *category spread* (r_p(133) = .2353, p = .0062) and *viewing photos* (r_p(133) = .2048, p = .0176) demonstrating that taking more positive photos, varying the categories in which they belong and viewing more old photos are associated with positive mood. Additionally, a more extended study could reveal a possible correlation with *photo entry consistency* as it is close to $p < 0.05$.

From these results we can conclude that the consistent usage of, and engagement with, this photo taking and annotation intervention is closely related with a positive change of the participant's psychological state. We can also conclude that the act of revisiting photographic content that participants generated to document positive events is significantly related with positive mood and affective change. However, we cannot reject the null hypothesis for the *average description duration, description* *edits*, *average description positivity* and *timeline explored* features as they did not display any significant correlations with the psychological scores.

4.4.3 Analysis per Day

The mood reports submitted by participants allowed us to explore potential correlations between the engagement of participants within the app and their daily mood. Participants were required to report their mood once at the end of each day and were cleaned using the mood-related exclusion criteria in Section 4.2.3. The reported mood values were standardised using z-score.

Following cleaning, the dataset consisted of 1812 individual days of mood reports across 70 participants (out of the 74 who completed the study). Our aim was to explore correlations between daily mood and the participant's activities. For each day we extracted features which represented the participant's activity within the application for that day. The features that were extracted are similar to those calculated for the survey period analysis, but in this case they were calculated per day. Therefore, the null hypothesis was that specific behavioural features (IV) will not correlate with the participant's mood (DV) during that day.

4.4.3.1 Results

We used Spearman's correlation coefficient to calculate the relationship between the mood reports and each feature. Spearman was chosen to minimise the effects of the outliers present in some of the data. The results are shown in Table 4.3.

The *photos taken* ($r_s(1811) = .0832$, $p = .0004$) and *category spread* ($r_s(1811) = .0903$, $p = .0001$) features display weak but significant correlations with mood, showing that logging more entries per day and spreading the categories in which they belong are correlated with a more positive mood. A participant who logs more entries is displaying a higher engagement with the intervention which has been shown to improve the effectiveness of such interventions [97]. *Category spread* is a feature

FEATURE	*Daily mood*
Photos taken	**0.0832*****
Category spread	**0.0903*****
Avg. desc. length	**0.0696*****
Avg. desc. duration	**0.0516***
Avg. desc. positivity	**0.0780***
Photos viewed	0.0460
Timeline explored	-0.0111

Table 4.3: Correlation results of daily mood features. $^{*}p <$ **.05**, $^{***}p <$ **.001**

we wanted to explore after the work by Parks et al. [125] discussed the positive impact of completing an assortment of tasks within a single intervention. *Description length* (r_s(1811) = .0696, p = .0031), *duration* (r_s(1811) = .0516, p = .0282) and *positivity* (r_s(1811) = .0780, p = .0146) show that the time and dedication one applies to the entry descriptions are also significantly correlated with mood. The *photos viewed* feature is also close to $p < 0.05$ and might show a significant correlation with more data considering that survey correlations with the same feature were found in Section 4.4.2. We can therefore conclude that taking more photos, engaging in a variety of different categories of positive experiences and writing longer and more positive descriptions are activities which are related to positive mood. However, we cannot reject the null hypothesis for the *photos viewed* and *timeline explored* features as they did not display significant correlations. Compared to the results from the survey periods, the correlation coefficients are relatively weaker. However, this is somewhat expected as the daily analysis considers short-term relationships within a single day.

4.4.4 Entry Categories

Correlating the number of entries belonging to the five different categories allowed us to observe whether any specific category was more positively correlated with psychological state.

As shown in Table 4.4, the positivity category shows significant correlations with

ΔP_{neg} (r_p(133) = -.2017, p = .0195) and $\bar{M}$ (r_p(1811) = .2149, p = .0126) and the reminiscence category shows significant correlations with $\bar{M}$ (r_p(1811) = .2338, p = .0066). None of the categories show significant correlations with ΔP_{pos} or ΔS and kindness, gratitude and destiny show no significant correlations with any of the psychological features. The reason for the correlations with positivity and reminiscence might be due to the unequal distribution of the number of entries in each category, shown in Figure 4.15. It might also be due to satisficing because those categories are at the top of the category selection screen (Figure 4.2) and require the least amount of cognitive effort to complete. Considering the lack of a distinct influential category across psychological features, we cannot conclude that any specific category has a larger impact on the intervention's efficacy. This might be a focus for future intervention research.

CATEGORY	ΔP_{pos}	ΔP_{neg}	ΔS	$\bar{M}$	*Daily mood*
Positivity	-0.0727	-0.0808	-0.0047	**0.2149***	0.0303
Reminiscence	-0.0121	**-0.2017***	0.0698	**0.2338****	-0.0440
Kindness	0.1225	-0.0668	-0.0040	0.0677	0.0310
Gratitude	-0.0130	-0.0162	-0.0145	0.0849	0.0144
Destiny	-0.0925	0.0395	-0.0136	0.0047	0.0284

Table 4.4: Correlation results of categories with psychological features. $^{*}p < .05$, $^{**}p < .01$

4.4.5 Image Analysis

During the study participants submitted 2178 photos with a range of content. As part of the analysis we wanted to explore if the type of content that participants generated is correlated with their change of psychological state. We utilised the IBM Watson's Visual Recognition deep learning service [62] to automatically extract information about the content of the photos. Watson extracted a range of descriptors about the content of each photo, including types of objects, number of faces and the dominant colour in each photo. Watson returned an array of classes, each with a score representing the service's confidence in the recognised class. Results could contain a class hierarchy like "/drink/beverage/alcoholic beverage/liquor/whiskey" representing different levels of detail for the recognised content.

Following the retrieval of results from Watson, we further cleaned up the image recognition dataset. We extracted the class with the highest confidence score which generally resulted in the least specific class in the hierarchy being extracted ("drink" in the previous example). Classes with a confidence score below 70% were discarded.

Watson detected a total of 221 unique classes across our dataset, of which the top 10 most common included: food (11%), indoors (10%), person (7%), building (6%), machine (5%), electronic device (3%), animal (3%), nature (3%), people (2%) and plant (2%). 23% of the photos were not classified, either due to a score below the threshold or Watson being unable to determine the contents of the photo. After extracting the classes we were left with many which were too specific, therefore we manually re-categorised them into 6 distinct groups: buildings/places (26%), objects (19%), food/drink (12%), people (10%), nature (7%) and other (4%). The number of faces in each photo was also extracted, with 20% of photos containing a single face and 17% containing two or more faces.

Watson also extracted the dominant colour in each photo. 50 unique colour classes were extracted by choosing the colour with the highest confidence score in each photo. These were then grouped into standard colours: red (16%), green (15%), black (15%), grey (15%), brown (11%), blue (8%), yellow (4%), pink (3%), orange (3%), purple (3%), white (2%) and unclassified (5%).

Figures 4.17 through 4.20 display some examples of the variety of photos taken by participants (cropped to fit). The photo in Figure 4.17 was first classified by IBM Watson as a "building" (90% accuracy) of "gray color" (77% accuracy) which was then manually grouped into "buildings/places" and "grey". This entry was categorised by the participant as "gratitude" who wrote about feeling angry but that going to church and praying and crying brought them relief. Figure 4.18 was first classified as "food" (96%) and "beige" (97%) which was then grouped into "food/drink" and "brown". This entry was also categorised as "gratitude" and described being grateful towards

Figure 4.17: A photo classified by IBM Watson as “buildings/places” and “grey”.

Figure 4.18: A photo classified by IBM Watson as “food/drink” and “brown”.

Figure 4.19: A photo classified by IBM Watson as “nature” and “blue”.

Figure 4.20: A photo classified by IBM Watson as “object” and “black”.

an acquaintance who bought them coffee and offered career advice. Figure 4.19 was first classified as “nature” (87%) of “blue color” (99%) which was then grouped into “nature” and “blue”. This participant was reminiscing about a prior holiday and aptly placed the entry into the “reminiscence” category. Lastly, Figure 4.20 was first classified as “keyboard” (79%) and “coal black” (97%) which was then grouped into “objects” and “black”. The participant categorised this entry as “positivity” and wrote about finally playing their keyboard.

First, we tried to see how the different classes and colours of content were related to the reported mood on the particular date they were captured. Each photo was annotated with the standardised mood ($range$ = -1–1) on that particular date. We then explored the distribution of mood reports for each class (Figure 4.21) and colour (Figure 4.22) of photo content. We can observe that for most classes and colours the distributions have averages close to 0 with negative skews. The distribution of photos classified as containing people showed a slightly more positive mood average of 0.04 (Mdn = 0.07) compared with the other classes, thus appearing to have a weak, positive relationship with mood.

A similar observation can be made for photos with a dominant colour of red (M = 0.04, Mdn = 0.07), orange (M = 0.03, Mdn = 0.06), brown (M = 0.02, Mdn = 0.08), black (M = 0.02, Mdn = 0.07) and grey (M = 0.01, Mdn = 0.07). This is an unexpected finding considering black, grey and brown are not usually regarded as "positive" colours [103]. Red and orange can have various cross-cultural meanings including warning and danger but also vibrant, emotional and pleasant [103] which might explain these relationships. These colours are then followed by blue, pink, white, yellow and green which are generally thought of as brighter colours. Interestingly the green photos, which could perhaps be associated with nature, also do not seem to have a very strong relationship with mood. There have been past papers conducting further analysis of an image's colours by extracting values for hue, saturation and brightness [146]; however, this work was specifically interested in Instagram filters. Because SnapAppy does not include any editing tools or filters for the photos, and participants were not specifically instructed to consider colour when taking photos, these findings would require a further, more targeted, study in order to determine whether there is a strong relationship between the colour of photos and mood.

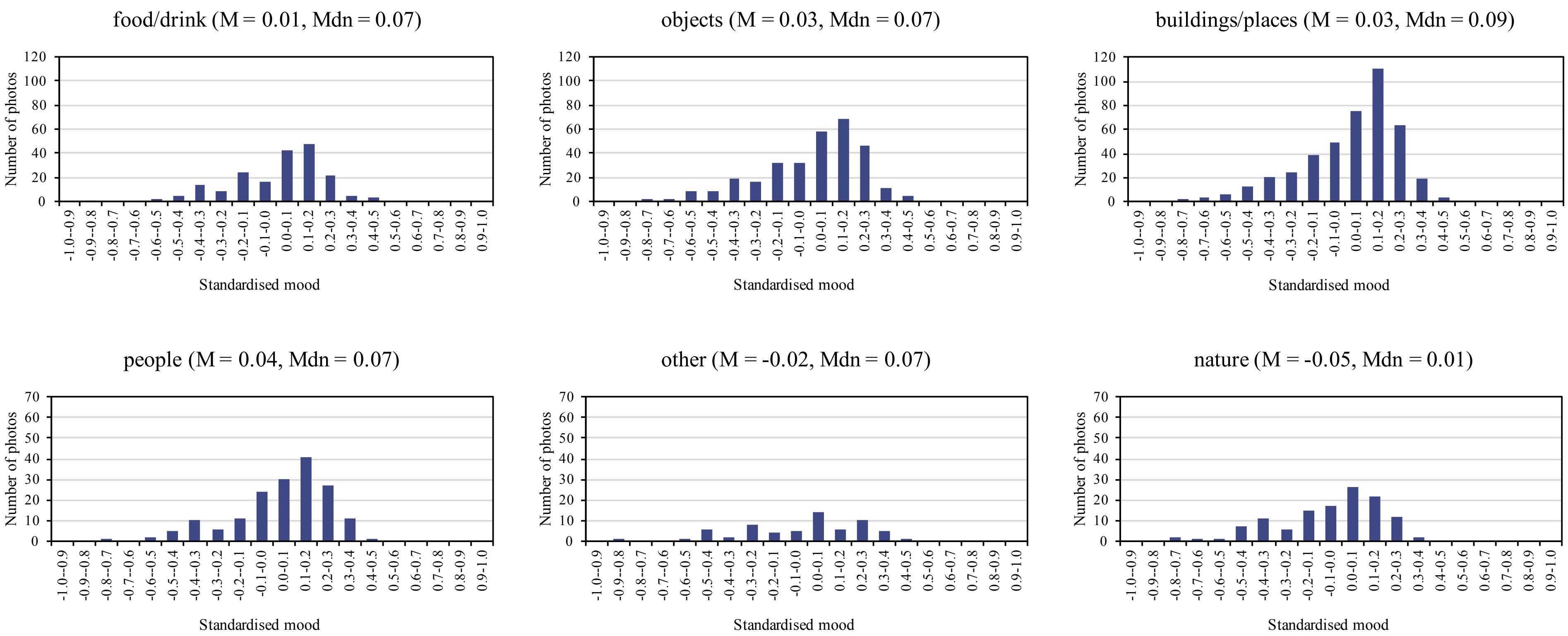

Figure 4.21: Distribution of photos by their associated standardised mood reports for each class group.

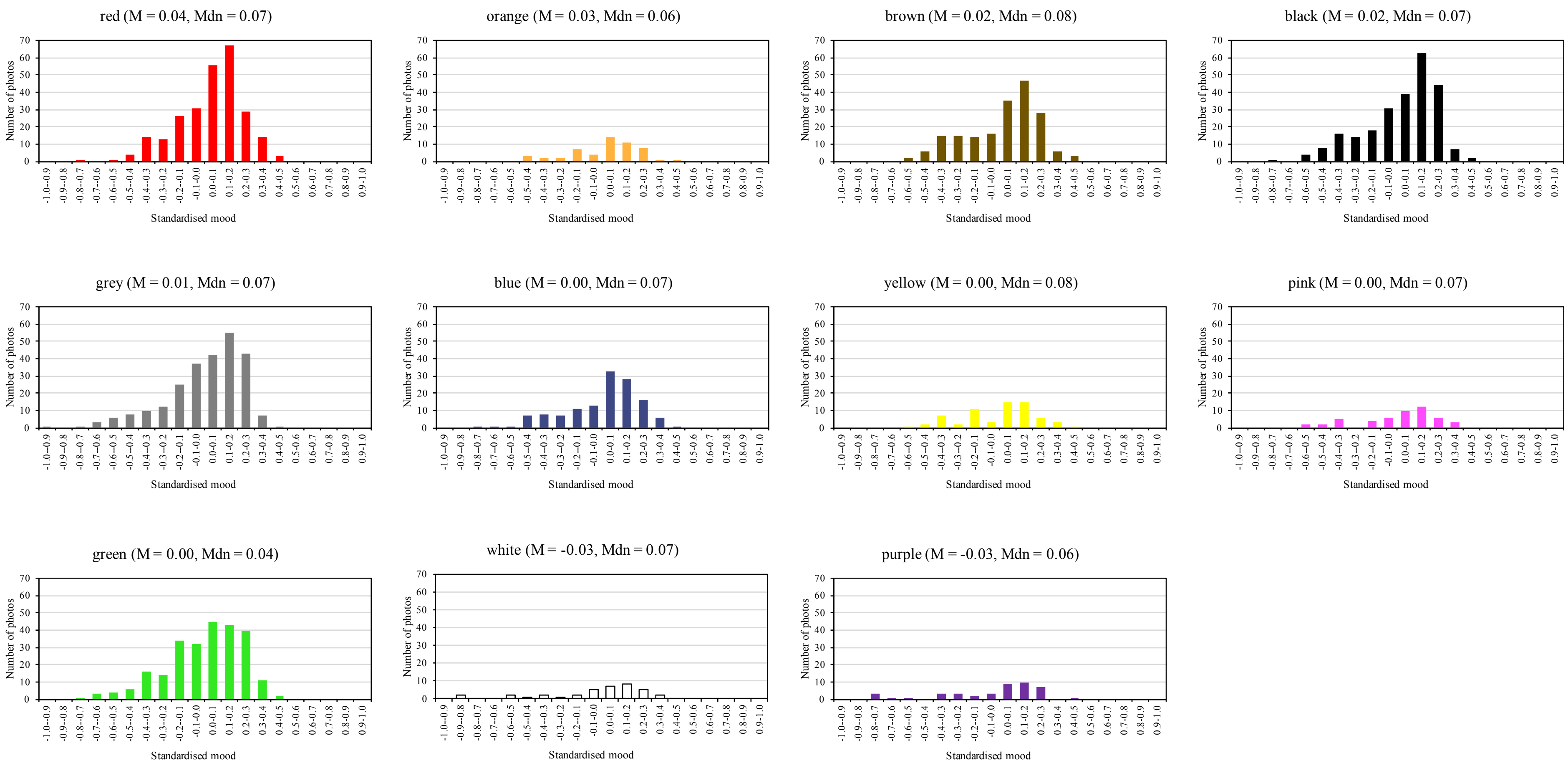

Figure 4.22: Distribution of photos by their associated standardised mood reports for each colour group.

Following these observations, we considered that the number of people present in a given photo could be a significant feature to explore. The correlations between the number of faces in the photos and the psychological changes over a survey period and the daily mood were calculated. Although the results from the survey period analysis did not show any significant correlations, we did observe a weak but significant correlation (r_s(1811) = .0684, p = .0034) between the number of faces appearing in photos and the daily mood reports (Table 4.5).

FEATURE	ΔP_{pos}	ΔP_{neg}	ΔS	$\bar{M}$	$Daily\ mood$
Avg. number of faces	-0.0075	-0.0119	0.0572	0.0418	**0.0684****

Table 4.5: Correlation results of no. of faces with psychological features. $^{**}p <$ **.01**

4.5 Key Findings

Following the analysis of the intervention's datasets, we are in a position to answer the research questions outlined in Section 4.1.

RQ3: Do smartphone photography activities correlate with changes in mood, affect and satisfaction with life?

Analysing the overall affect of the intervention, the Wilcoxon signed-rank test revealed a significant improvement in satisfaction with life between the start and end of the study. However, after observing the raw scores in Figure 4.16 it was clear that, although the average SWLS scores increased over the duration of the study, there were a wide range of scores across the intervention. This demonstrated that simply joining the study would not necessarily lead to a clear improvement in the participant which motivated the analysis of specific photography activities. Based on the analysis of results we can see significant correlations between the level of engagement of participants with smartphone photography as a form of intervention and their positive change of affective state over $\sim$10 day periods and mood changes on a daily basis.

RQ4: Which specific smartphone photography activities show a statistically significant correlation with change in mood, affect and satisfaction with life?

In our analysis we have identified a number of photo taking activity patterns that are related to positive change in mood and affect. The length of photo annotations is associated with a positive change in PANAS scores. Frequent and consistent photo taking over the duration of the study, spread across multiple categories is significantly correlated with negative affect and positive mood. Reviewing past photos is related to a reduction in negative PANAS scores and positively correlated with reported mood. Finally, with respect to the photo content, the number of faces in the photos shows a weak, positive correlation with mood reports.

It is important to emphasise the significance and limitations of the "intervention bias" and "selection bias" that are commonly observed in such studies. These refer to the fact that the participants are aware that they are involved in a positive psychology intervention, and that those who chose to participate may have done so because they experience negative emotional well-being and want to improve it. The intervention bias is a key aspect of most psychological interventions and in many cases it is a desirable effect that is part of the intervention mechanism itself. In most interventions, the desired effects can be achieved only when the participants are fully aware of the purpose of the intervention. We did not specifically study this phenomenon in this chapter but it was incorporated into the study and app design. Regarding selection bias, our data showed that there was such effect in our study as well. The average PANAS scores when participants joined the study were slightly lower than the average scores for the general population. Our results show that for such individuals, engaging with smartphone photography as a form of positive psychological intervention can be associated with positive changes in their mood and affective state.

4.6 Discussion

Photo taking and viewing has always played an emotional role in people's lives. Traditional photos and tangible photo albums were always objects with significant emotional effect, either as ways of capturing happy (or unhappy) memories or by invoking positive feelings when we revisit them. The proliferation of smartphone photography has offered new opportunities for people to easily capture a range of moments in their lives that would have been significantly more difficult in the past. However, the relative ease of smartphone photography has also enabled a more "relaxed" attitude towards photo taking and reviewing. Users may take photos at a significantly higher rate than before with less consideration about the meaning or importance of the subject of the photos. This random nature of spontaneous smartphone photography can lead to a more disorganised photo taking practice and possibly a lower rate of revisiting older photos mainly due to the significantly increased quantity.

Based on the findings of this study, we believe that there is an opportunity for amateur smartphone photography to play a more active role in people's affective state, if applied in an appropriate manner. For smartphone photography to act as a positive intervention there is a need for a more disciplined and structured engagement with photo taking. In this study, participants were expected to take photos that had a particular meaning which was directly related to their psychological state. Regular taking of such photos and reviewing is shown to be related to positive changes in mood and affect. Considering existing photo taking systems, either stand-alone photo galleries, or social network integrated applications (e.g. Instagram or Facebook), the explicit indication of purpose related to mood or emotion would be key. The possible integration of "emotional purpose" annotation on photos (possibly related to the categories used in this study) can play that role. This technique would enable users to consider the emotional content of their photos and the role they may play in helping them re-experience particular emotions. Furthermore, such explicit annotation can instil a more focused and engaged attitude towards emotionally meaningful photo taking.

Regarding reviewing of photos with potential emotional impact, there have already been examples developed by online services, such as Facebook's "Memories" feature (previously called "On This Day"), which prompts users to view their activity and posts from previous years. Such features can take a form of a more explicit intervention if they can be activated when the target user is in need of an emotional boost. As shown in the observations, lower PANAS scores were more prevalent suggesting a potential selection bias of participants who are more willing to engage with an intervention that can potentially improve their psychological state. The work in Chapter 3 demonstrated techniques to passively detect changes in mood by analysing the user's online activity [86] which could then trigger appropriate photo viewing opportunities when mood is low.

Our results display the potential to infer the emotional state of a user by analysing their photo taking behaviour. As shown in our study, certain behaviours of participants regarding photo taking and viewing are shown to be significantly correlated with their mood and affective state. This implies that a typical photo management application can potentially infer the emotional state of the user by analysing their photo taking and viewing patterns. Such a technique can act as a contextual input for a recommendation system, adapting a system's functionality according to the user's affective state.

Lastly, we must acknowledge the privacy concerns and ethical grey area around taking photos of others for this study. Participants were not instructed to take photos of any particular subject matter; however, many did take photos of their friends and family. Participants were fully aware and informed in detail of how their photos would be handled during the study. In particular, they were aware that their photos would only be available to the research team. The grey area applies to the app's design and the legality of photography. Outside of the research study, SnapAppy is of a similar nature to a private diary, such that photographed subjects may be more comfortable with their photos being kept within the app, compared to being shared on an online social network feed, for example. The legality of photography of people on public and private property varies by jurisdiction, location, subject, context and behaviour and the laws can often be vague or unclear. It is the responsibility of the photographer to understand their rights and to ensure that they are abiding by the law.

4.7 Future Work

The research in this study can be expanded in the future in terms of thematic and qualitative analysis.

In our analysis, we performed sentiment analysis on the photo descriptions in order to extract the positive or negative sentiment of the text using the LIWC software. A deeper analysis using the remaining linguistic, social, emotional and topical variables extracted by LIWC might be used to draw some additional conclusions about how the participants annotated their photos. In addition, thematic analysis might also be used to identify themes of discussion in the descriptions and to correlate those themes with changes in the participants' emotional well-being. This knowledge could be used to tailor future interventions involving writing tasks.

A post-study interview or questionnaire may have also provided extra insight into the participants' experiences with the study and the usability of the app, how their photo taking habits changed during the study, how their subjective well-being changed and how engaged they felt throughout the study, for example. Interview answers may able to able to provide some further insight into the 20-day intervention "sweet spot" discussed in Section 4.4.1. Personality and photography-proficiency questionnaires may have also revealed some interesting differences between particular participants and their behaviours in the study. Additionally, as with other positive psychology interventions, the positive effects of the study can subside over time. Conducting longitudinal emotional well-being surveys post-study might allow researchers to estimate how long the effect lasts.

Lastly, in addition to the discussion regarding a more structured engagement with photo taking, a contrasting study tracking participants' everyday photo taking behaviour without an explicit intervention, categories or guidance may yield interesting results when coupled with ecological momentary assessments to track emotional well-being. This may help to provide a better understanding of how non-therapeutic photography affects people in a real-world setting.

4.8 Conclusions

In this chapter we report on the results of a positive psychology intervention based on smartphone photography as a method to improve mood and affect. As discussed in Section 2.4.2, the limited state-of-the-art literature in this field focuses on the use of photography in counselling (phototherapy) [52], mindful and creative photo taking [80], a guided photo taking intervention including selfies and photos to induce happiness [20] and technology mediated reflection utilising photos [65]. This novel study focuses on the integration of smartphone photography and positive psychology practices. The methodology design is highly motivated by the well established positive psychology intervention literature, aiming to translate the techniques revolving around writing tasks into a photography-based intervention. Participants were encouraged to take photos about positive moments, events and experiences and revisit those photos using the SnapAppy app on their smartphone. The traditional psychological interventions were translated into the five photo categories which were used to guide the participants to take positive photos. We evaluated the effectiveness of the method through common psychological surveys including PANAS, SWLS and daily mood reports. Our analysis also displays novelty when compared with many other positive psychology intervention studies. Whereas commonly only pre-test and post-test scores are compared, our design follows the ecological momentary assessment methodology by assessing participants throughout the study in order to observe the changes on a more granular level. Our results show that regular taking of positive photos, frequent reviewing of photos and taking pictures of people, have positive correlations with the change of the participant's mood and affective state. We hope that the positive results of this study will motivate a wider exploration of smartphone photography as a potential instrument to improve people's emotional well-being.

— 5 —

Conclusions

5.1 Thesis Summary and Contributions

In the preceding chapters I have demonstrated and discussed the results of my research, presenting two novel case studies exploring the passive detection of mood using behavioural data from online social networks and utilising smartphone photography as a form of positive psychology intervention. These results and contributions build upon the existing literature discussed in Chapter 2 and are summarised below.

The literature review in Chapter 2 explored the various aspects of emotional well-being including affective state, emotion, mood, wellness and mental health conditions. Next, different forms of experience sampling methodologies were discussed, followed by the traditional and digital assessments for capturing these affective phenomena. The challenges and limitations of manual assessment methodologies motivated the review of literature regarding the passive detection of different aspects of emotional well-being. Specifically, smartphone sensing, wearable sensing and online behavioural tracking methods were discussed. This was followed by a discussion of different types of positive psychology intervention, specifically traditional life review interventions and more modern technology-enhanced interventions, which were used to improve the participants' emotional well-being. The challenges associated with positive psychology interventions are then outlined. Lastly, the literature review covers real-world applications of tracking and influencing emotional well-being, including lifelogging, visualisation and affective computing.

Chapter 3 presented a study exploring the use of online social network activity as an indicator of mood changes. During a month-long study, online behavioural data from Facebook and Twitter was captured and real-world mood reports as ground truth were collected via experience sampling using a smartphone application. The collection of data from multiple social networks and the use of experience sampled mood reports was a novel approach in comparison to past literature which would rely on data from a single social network providing an accurate, truthful representation of one's emotional well-being. The data was used in order to build a machine learning classifier to passively infer the participants' real-world mood changes. The results contributed to the following research questions:

RQ1: Which specific behavioural activities on Facebook and Twitter correlate with real-world mood changes?

During analysis we extracted a set of features which represented the participant's online behaviour across both social networks including the number of likes, comments, tweets, mentions and retweets and the time of each post. We also extracted some calculated features such as text sentiment, the total activity across the social networks and active (posts, comments, tweets, replies) and passive (likes, retweets) activities. Calculating the Pearson's correlation for each feature and the participant's mood changes across a 7-day sliding window allowed us to observe the features which showed statistically significant results for the most participants. As shown in Figure 3.3, the total, combined activity across both Facebook and Twitter is the top correlated feature, followed closely by Facebook activity alone, the number of days the participant was last active on Facebook and their passive and active activities. This result emphasises the motivation of this paper, which discussed how relying on a single online social network as an accurate representation of one's psychological state may not be reliable. The features combining data from both Facebook and Twitter show a higher number of participants with significant correlations. These results are not only valuable to future research specifically focusing on Facebook and Twitter; the extracted features are also transferable to research into additional social networks.

RQ2: Can activity from both Facebook and Twitter be used to infer real-world mood changes?

As shown in Figures 3.5 and 3.6, it was found that the correlation between online activity and mood changes was positive for some participants and negative for others. This may be due to the reasons why certain people use online social networks and the type of content they are observing and posting. Although there is anecdotal evidence that links particular moods with online behaviour, the aim of this work is not to identify an explicit causal relationship. Instead we attempted to classify users into two groups: strong and weak correlations. Then those who portrayed strong correlations would be classified into positive or negative groups (see Figure 3.1). A set of features for use in the classifier were identified including the length of the textual posts, the ratio of active and passive activities and the ratio of Facebook and Twitter activities. These features conceptually captured the level of commitment by the participants when interacting with the OSNs. As shown in Table 3.3, the "Strong vs Weak" classifier demonstrated 95.2% precision and 94.7% recall and the "Positive vs Negative" classifier demonstrated 84.4% precision and 80.0% recall. This result shows that online social network activity data from Facebook and Twitter can be used to infer mood changes over a 7-day sliding window with relatively high accuracy. This contribution furthers the existing approaches to enable the passive tracking of emotional well-being without participant input or burden, improving many real-world applications including life-logging, affective computing, counselling and mental health care. This research will also contribute to the awareness of verifying inferred emotional states with real-world data.

Chapter 4 presented a positive psychology intervention involving momentary smartphone photography as a method of improving emotional well-being. This novel research fills the gap in the integration of smartphone photography and positive psychology by analysing the positive consequences of taking and reviewing photos related to positive experiences. The study saw 74 participants use a smartphone application called SnapAppy for one month, taking and reviewing photos related to positive moments, events and experiences. Participants completed daily mood reports alongside PANAS and SWLS surveys every 10 days. The analysis identified the results to the following research questions:

RQ3: Do smartphone photography activities correlate with changes in mood, affect and satisfaction with life?

The initial stage of the analysis was to observe any overall changes from participating in the study. A Wilcoxon signed-rank test was used to explore if there were any significant changes between the participant's surveys at the beginning and end of the study. The results did not show any significant changes in the positive and negative PANAS scores but did show a significant change in the SWLS scores. However, after observing the scores more closely in Figure 4.16, it was clear that there was a significant range of scores across the intervention, which demonstrated that participating in the study would not necessarily lead to a clear improvement in positive and negative affect or satisfaction with life. These results then lead into the following research question which aims to analyse each activity within the intervention more deeply in order to discover specific correlations between smartphone photography and improvements in mood, affect and satisfaction with life.

RQ4: Which specific smartphone photography activities show a statistically significant correlation with change in mood, affect and satisfaction with life?

The SnapAppy app contained a comprehensive logging system which tracked every action the participants performed within the app when taking part in the study. From this data we were able to extract a set of features which are described in

Table 4.1. Surveys were completed approximately every 10 days and thus formed a convenient period in which to divide the data. The activity within each period was correlated with the change in PANAS and SWLS scores across the period and the average mood during the period to observe any improvements to mood, affect and satisfaction with life. A similar analysis methodology was also performed for the daily mood reports. The correlation results showed that the length of photo annotations was associated with a positive change in PANAS scores. Frequent and consistent photo taking over the duration of the study, spread across multiple categories was significantly correlated with negative affect and positive mood. Reviewing past photos was correlated with a reduction in negative PANAS scores and positively correlated with mood. In addition to the contribution of a novel method of positive psychology intervention, the extracted features representing the participants' behaviours are transferable to future interventions involving photo taking and reviewing and also to existing photo taking systems as discussed in Section 4.6.

In addition to the analysis of behavioural features, we also examined the effect of the different entry categories and image features extracted from the photos taken by the participants. The goal was to observe whether any of the five entry categories was more influential to the participants' psychological state. Although some significant correlations were present for the positivity and reminiscence categories, it was concluded that the lack of a distinct influential category across psychological features meant that we could not define any specific category as having a larger impact on the intervention's efficacy. Secondly, IBM Watson's Visual Recognition deep learning service was used to extract the objects, faces and dominant colours in each photo. After grouping the object classifications and colours, we attempted to observe how they were related to the reported moods on dates the photos were captured. We observed that for most classes and colours the distributions had averages close to 0 with negative skews. Photos containing people, red and orange showed slightly more positive mood averages. This result prompted us to calculate the correlations of the number of faces in each photo with the psychological features which revealed a weak but significant positive correlation with daily mood.

Together, the methodologies and technologies presented in these studies can work together to provide a pragmatic system to improve emotional well-being. Psychological assessments, active tracking, passive detection and interventions form an important relationship to first develop the awareness of one's emotional well-being, either through self-awareness or passively via machine learning. Secondly, that information can then be used as motivation or an input for an intervention to make positive changes to one's life. Without the intervention to make changes, collecting and analysing the data would be meaningless. Without the data as input, the intervention can be ungrounded and unguided. Thus our contributions provide novel advancements to the system as a whole.

5.2 Limitations

In addition to the challenges and limitations discussed in the previous chapters, there were a set of general limitations that were experienced across the work presented in this thesis.

As with any research study, we strive to collect as much data as possible in order to improve the reliability of our findings. As previously discussed, participant compliance and dropout rates have the tendency to worsen for longitudinal research as participants experience fatigue towards the study. When designing the one-month long studies presented in this thesis, we aimed to ensure the tasks required of the participants were as easy to complete as possible in order to retain compliance. The final datasets used for analysis consisted of 16 participants for Chapter 3 and 74 for Chapter 4. Although 16 participants may be considered too few, the dataset for each participant consisted of one-month's worth of data, thus ensuring that the quantity of data was sufficient to draw interesting and novel conclusions. Nevertheless, running a large scale study with more participants may have resulted in even stronger or more interesting findings. A larger dataset would also provide a greater spread of ages and genders which would allow for a deeper analysis into how different demographics might display differing behaviours.

As discussed in the methodology sections, the psychological assessment methods were chosen with careful deliberation after consultations with psychologists and reviewing the relevant literature to ensure that they were suited for each study's purpose. As is the case with many longitudinal studies, participant fatigue can become problematic if not managed correctly. We specifically identified assessments which were both commonly used in the literature and short in completion time to aid in the reduction of burden and fatigue of our participants. This also allowed us to conduct assessments throughout the studies rather than only before and after. The limitation of this decision was that we were unable to conduct any long-form assessments such as personality tests or assessments for several specific emotions which may have revealed some interesting findings.

Also in relation to participants, the selection bias phenomenon has the potential to have an impact on the type of research covered in this thesis. In Chapter 3, the inclusion criteria required participants who were active users of both Facebook and Twitter. This criteria can limit the demographics of participants to those who are active social network users and to cultures, countries and communities where Facebook and Twitter are popular. However, the highly correlated features discovered in the analysis could be transferable to research with other social networks. The selection and intervention bias is also likely to have had an impact on the work in Chapter 4. People with mental health issues may be actively searching for self-help solutions and may be more likely to enrol in a positive psychology intervention study. Although this is not necessarily a negative result, it ensures that the dataset will not be a well-rounded representation of the population. With a large scale study, a component of the recruitment process could be to pre-screen participants in order to ensure a well-rounded sample.

Finally, any application which actively or passively tracks and logs user activity and behaviour raises privacy concerns. This can impact recruitment for research studies; participants may not feel comfortable sharing their personal data even when the institution is trusted and incentives are provided. For those who *do* participate, the lingering privacy concerns may cause participants to alter their behaviour from their norm, which might impact the validity of the results. Ideally the studies could be conducted without the participants' prior knowledge of being tracked in order to obtain their true normal behaviour; however, this is ethically questionable, especially when the data involves personal social networks and mental health.

5.3 Future Work

The research and results presented in this thesis raise some compelling directions for future work in the area of tracking and influencing emotional well-being.

The work in Chapter 3 reveals that activity across both Facebook and Twitter is a good indicator of mood changes; however, this was discussed to be limited to those who are active users of both social networks. Future research across different sets of social networks, other than Facebook and Twitter, would ensure that the methodology and extracted features are transferable across to different OSNs. It would also ensure this research is more globally applicable in countries and communities where Facebook and Twitter are not as popular. The features are pertinent to social networks which are focused on textual posts, social reactions and comments; however, other platform specific features could also be included such as tags or social ties.

The results in this work were achieved by analysing mood changes over a 7-day sliding window. Reducing the window duration will allow for a system which can predict mood changes more accurately over a shorter time period. A future study collecting data from more participants with more frequent mood reports during a day, or different types of psychological assessments, may help to achieve this goal.

These suggestions for future work will contribute to a more accurate passive mood tracking system which could be applied to many academic and real-world scenarios. Future research into the psychological effects of social networks would be able to passively infer mood changes without inducing burden to the participants. Additionally, research utilising lifelogging or experience sampling of mood would only require passive access to the participants' social networks. Passive logging of mood would also have benefits in the self-help community and in mental health therapy, again reducing the required active participation by the subjects. Passive detection of negative changes in mood might also be used as a trigger in positive psychology interventions.

During the intervention presented in Chapter 4, participants could take photos and assign them to one of five categories: positivity, reminiscence, kindness, gratitude and destiny. In the analysis, we observed the correlations of the five categories with the participant's change in mood and affect. Replicating the results from previous literature, the spread of categories showed correlations with mood [125]; however, we did not come to any strong conclusions about the efficacy of each category. The implications of these specific tasks and the optimal number of different tasks might be considered for future research.

Research into the conscious and subconscious relationship between colour and psychological state is an interesting area which could be contributed to through a future study involving momentary smartphone photography. Existing literature already analyses the subconscious use of Instagram filters to predict markers of depression [146]; however, a more controlled focus could be placed on the use of colours to add conscious emotional meaning to photos taken by participants.

Although this process is often utilised by professional photographers in order to portray or induce certain emotions in their photos, this could be used in amateur photography situations such as phototherapy, allowing patients to easily convey their emotions through photographs.

Lastly, the correlation between a positive change in mood and affect and the number of times the participants viewed old photos was displayed in the results. This behaviour is encouraged in applications like Timehop [184], Facebook's "Memories" feature (previously called "On This Day") and the "Rediscover this day" feature within Google Photos. These are modern-day, technological versions of revisiting old printed photo albums with the purpose of inducing positive emotions in the viewer. For many, one of the purposes of sharing content online is to receive recognition from one's peers, usually in the form of likes, comments and reposts. This behaviour was defined as the "capitalization of positive emotions" where the process of publicly sharing positive experiences on OSNs enriches the derived emotional benefit beyond the original experience [160]. A compelling direction for future work could be to examine the psychological effect of revisiting posts with high amounts of recognition. For example, exploring whether revisiting photos with many likes and comments has a stronger effect compared to revisiting photos with fewer responses. If this is indeed the case, this phenomenon could be exploited as a positive psychology intervention to promote positive feelings.

5.4 Concluding Remarks

Positive emotional well-being is one of the fundamental components of a good quality of life and human flourishing. A decrease in emotional well-being can lead to serious mental and physical health conditions, which much of the literature discussed in this thesis aims to combat. The work in this thesis presented two novel case studies focused on exploring the use of online social network activity to passively track mood and smartphone photography as a positive psychology intervention to influence mood and affect. I hope that these contributions will influence and inspire future research into tracking and influencing emotional well-being. Contributing to the happiness and positivity of even a single person is considered a success.

Bibliography

[1] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, “Sentiment analysis of Twitter data,” in *Proceedings of the Workshop on Languages in Social Media*. Association for Computational Linguistics, 2011, pp. 30–38.

[2] N. Andalibi, P. Öztürk, and A. Forte, “Sensitive self-disclosures, responses, and social support on Instagram: The case of #depression,” in *Conference on Computer Supported Cooperative Work and Social Computing*, 2017, pp. 1485–1500.

[3] Apache, “Apache Cordova,” https://cordova.apache.org/, 2018, accessed: 2018-01-29.

[4] M. Argyle, “Non-verbal communication in human social interaction,” in *Non-Verbal Communication*. Cambridge U. Press, 1972, pp. 243–270.

[5] S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias, “Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment,” *Multimedia Tools and Applications*, vol. 41, no. 3, pp. 469–493, 2009.

[6] M. D. Back, J. M. Stopfer, S. Vazire, S. Gaddis, S. C. Schmukle, B. Egloff, and S. D. Gosling, “Facebook profiles reflect actual personality, not self-idealization,” *Psychological Science*, vol. 21, no. 3, pp. 372–374, 2010.

[7] R. C. Balabantaray, M. Mohammad, and N. Sharma, “Multi-class Twitter emotion classification: A new approach,” *International Journal of Applied Information Systems*, vol. 4, no. 1, pp. 48–53, 2012.

[8] J. B. Bayer, N. B. Ellison, S. Y. Schoenebeck, and E. B. Falk, “Sharing the small moments: Ephemeral social interaction on Snapchat,” *Information, Communication & Society*, vol. 19, no. 7, pp. 956–977, 2016.

[9] BBC, “Facebook-Cambridge Analytica data scandal,” https://www.bbc.co.uk/news/topics/c81zyn0888lt/facebook-cambridge-analytica-data-scandal, 2018, accessed: 2018-08-24.

[10] A. T. Beck and R. A. Steer, “Manual for the Beck anxiety inventory,” *The Psychological Corporation*, 1990.

[11] A. T. Beck, R. A. Steer, and G. K. Brown, “Beck depression inventory-ii,” *The Psychological Corporation*, vol. 78, no. 2, pp. 490–8, 1996.

[12] J. S. Beck, *Cognitive behavior therapy: Basics and beyond*. Guilford press, 2011.

[13] A. Betella and P. F. Verschure, “The affective slider: A digital self-assessment scale for the measurement of human emotions,” *PLOS ONE*, vol. 11, no. 2, p. e0148037, 2016.

[14] Blinky, “Perspective a new journaling experience,” http://blinky.co/perspective_app/, 2017, accessed: 2018-08-13.

[15] N. Bolger, A. Davis, and E. Rafaeli, “Diary methods: Capturing life as it is lived,” *Annual Review of Psychology*, vol. 54, no. 1, pp. 579–616, 2003.

[16] J. Bollen, H. Mao, and A. Pepe, “Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena,” *International Conference on Weblogs and Social Media*, vol. 11, pp. 450–453, 2011.

[17] M. M. Bradley and P. J. Lang, “Measuring emotion: The self-assessment manikin and the semantic differential,” *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[18] C. M. Burton and L. A. King, “The health benefits of writing about intensely positive experiences,” *Journal of Research in Personality*, vol. 38, no. 2, pp. 150–163, 2004.

[19] L. Canzian and M. Musolesi, “Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis,” in *International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 1293–1304.

[20] Y. Chen, G. Mark, and S. Ali, “Promoting positive affect through smartphone photography,” *Psychology of Well-Being*, vol. 6, no. 1, p. 8, 2016.

[21] Y. Chen, G. Mark, S. Ali, and X. Ma, “Unpacking happiness: Lessons from smartphone photography among college students,” in *International Conference on Healthcare Informatics*. IEEE, 2017, pp. 429–438.

[22] E. K. Choe, N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz, “Understanding quantified-selfers’ practices in collecting and exploring personal data,” in *Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 1143–1152.

[23] P. Cipresso, S. Serino, D. Villani, C. Repetto, L. Sellitti, G. Albani, A. Mauro, A. Gaggioli, and G. Riva, “Is your phone so smart to affect your state? An exploratory study based on psychophysiological measures,” *Neurocomputing*, vol. 84, pp. 23–30, 2012.

[24] E. A. Cook, “Effects of reminiscence on life satisfaction of elderly female nursing home residents,” *Health Care for Women International*, vol. 19, no. 2, pp. 109–118, 1998.

[25] J. R. Crawford and J. D. Henry, “The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample,” *British Journal of Clinical Psychology*, vol. 43, no. 3, pp. 245–265, 2004.

[26] J. D. Creswell, “Mindfulness interventions,” *Annual Review of Psychology*, vol. 68, pp. 491–516, 2017.

[27] M. C. Davis, “Life review therapy as an intervention to manage depression and enhance life satisfaction in individuals with right hemisphere cerebral

vascular accidents," *Issues in Mental Health Nursing*, vol. 25, no. 5, pp. 503–515, 2004.

[28] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media," *International Conference on Weblogs and Social Media*, vol. 13, pp. 1–10, 2013.

[29] M. De Choudhury, S. Counts, E. J. Horvitz, and A. Hoff, "Characterizing and predicting postpartum depression from shared Facebook data," in *Conference on Computer Supported Cooperative work & Social Computing*. ACM, 2014, pp. 626–638.

[30] M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar, "Discovering shifts to suicidal ideation from mental health content in social media," in *Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 2098–2110.

[31] A. de Santos Sierra, C. S. Ávila, J. G. Casanova, and G. B. del Pozo, "A stress-detection system based on physiological signals and fuzzy logic," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4857–4865, 2011.

[32] DefenceHealthAgencyConnectedHealth, "T2 Mood Tracker," http://t2health.dcoe.mil/apps/t2-mood-tracker, 2016, accessed: 2018-07-30.

[33] L. Dennison, L. Morrison, G. Conway, and L. Yardley, "Opportunities and challenges for smartphone applications in supporting health behavior change: Qualitative study," *Journal of Medical Internet Research*, vol. 15, no. 4, 2013.

[34] F. G. Deters and M. R. Mehl, "Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment," *Social Psychological and Personality Science*, vol. 4, no. 5, pp. 579–586, 2013.

[35] E. Diener, "Subjective well-being," *Psychological Bulletin*, vol. 95, no. 3, p. 542, 1984.

[36] E. Diener, R. A. Emmons, R. J. Larsen, and S. Griffin, "The satisfaction with life scale," *Journal of Personality Assessment*, vol. 49, no. 1, pp. 71–75, 1985.

[37] E. Diener, D. Wirtz, W. Tov, C. Kim-Prieto, D.-W. Choi, S. Oishi, and R. Biswas-Diener, "New well-being measures: Short scales to assess flourishing and positive and negative feelings," *Social Indicators Research*, vol. 97, no. 2, pp. 143–156, 2010.

[38] B. R. Duffy, "Anthropomorphism and the social robot," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 177–190, 2003.

[39] M. Eid and R. J. Larsen, *The science of subjective well-being*. Guilford Press, 2008.

[40] P. Ekkekakis, *The measurement of affect, mood, and emotion: A guide for health-behavioral research*. Cambridge University Press, 2013.

[41] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[42] P. Ekman, W. V. Friesen, and P. Ellsworth, *Emotion in the human face: Guidelines for research and an integration of findings*. Elsevier, 2013.

[43] R. A. Emmons and M. E. McCullough, "Counting blessings versus burdens: An experimental investigation of gratitude and subjective well-being in daily life," *Journal of Personality and Social Psychology*, vol. 84, no. 2, p. 377, 2003.

[44] A. Exler, A. Schankin, C. Klebsattel, and M. Beigl, "A wearable system for mood assessment considering smartphone features and data from mobile ecgs," in *International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 2016, pp. 1153–1161.

[45] J. Fardouly and L. R. Vartanian, "Negative comparisons about one's appearance mediate the relationship between Facebook usage and body image concerns," *Body Image*, vol. 12, pp. 82–88, 2015.

[46] A. Fessl, V. Rivera-Pelayo, V. Pammer, and S. Braun, "Mood tracking in virtual meetings," in *European Conference on Technology Enhanced Learning*. Springer, 2012, pp. 377–382.

[47] F. Franchignoni, L. Tesio, M. Ottonello, and E. Benevolo, "Life satisfaction index," *American Journal of Physical Medicine & Rehabilitation*, vol. 78, pp. 509–515, 1999.

[48] C. A. Frantzidis, C. Bratsas, M. A. Klados, E. Konstantinidis, C. D. Lithari, A. B. Vivas, C. L. Papadelis, E. Kaldoudi, C. Pappas, and P. D. Bamidis, "On the classification of emotional biosignals evoked while viewing affective pictures: An integrated data-mining-based approach for healthcare applications," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 309–318, 2010.

[49] B. L. Fredrickson and M. F. Losada, "Positive affect and the complex dynamics of human flourishing," *American Psychologist*, vol. 60, no. 7, p. 678, 2005.

[50] J. H. Geer, "The development of a scale to measure fear," *Behaviour Research and Therapy*, vol. 3, no. 1, pp. 45–53, 1965.

[51] S. Gillet, P. Schmitz, and O. Mitas, "The snap-happy tourist: The effects of photographing behavior on tourists' happiness," *Journal of Hospitality & Tourism Research*, vol. 40, no. 1, pp. 37–57, 2016.

[52] M. M. Ginicola, C. Smith, and J. Trzaska, "Counseling through images: Using photography to guide the counseling process and achieve treatment goals," *Journal of Creativity in Mental Health*, vol. 7, no. 4, pp. 310–329, 2012.

[53] Google, "PACO," https://www.pacoapp.com/, 2018, accessed: 2018-08-20.

[54] M. C. Graham, L. Priddy, and S. Graham, *Facts of Life: Ten issues of contentment*. Outskirts Press, 2014.

[55] J. Greenwood, "MoodPanda rate and track your mood online," http://moodpanda.com/, 2018, accessed: 2018-08-13.

[56] R. Gummadi, "Instagram Graph API launches and Instagram API platform deprecation," https://developers.facebook.com/blog/post/2018/01/30/instagram-graph-api-updates/, 2018, accessed: 2018-07-19.

[57] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using bio-sensors: First steps towards an automatic system," in *Tutorial and Research Workshop on Affective Dialogue Systems*. Springer, 2004, pp. 36–48.

[58] G. M. Harari, S. R. Müller, V. Mishra, R. Wang, A. T. Campbell, P. J. Rentfrow, and S. D. Gosling, "An evaluation of students' interest in and compliance with self-tracking methods: Recommendations for incentives based on three smartphone sensing studies," *Social Psychological and Personality Science*, vol. 8, no. 5, pp. 479–492, 2017.

[59] M. Hasan, E. Rundensteiner, X. Kong, and E. Agu, "Using social sensing to discover trends in public emotion," in *International Conference on Semantic Computing*. IEEE, 2017, pp. 172–179.

[60] E. Hauthal and D. Burghardt, "Extraction of location-based emotions from photo platforms," in *Progress in Location-Based Services*. Springer, 2013, pp. 3–28.

[61] A. Howells, I. Ivtzan, and F. J. Eiroa-Orosa, "Putting the 'app' in happiness: A randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing," *Journal of Happiness Studies*, vol. 17, no. 1, pp. 163–185, 2016.

[62] IBM, "IBM Watson visual recognition," https://www.ibm.com/watson/services/visual-recognition/, 2018, accessed: 2018-01-12.

[63] InexikaInc, "iMoodJournal mood tracking mobile application," https://www.imoodjournal.com/, 2018, accessed: 2018-08-13.

[64] inSightOne, "MoodRunner," https://itunes.apple.com/us/app/mood-runner/id404822869, 2010, accessed: 2018-06-11.

[65] E. Isaacs, A. Konrad, A. Walendowski, T. Lennig, V. Hollis, and S. Whittaker, "Echoes from the past: How technology mediated reflection improves well-being," in *Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 1071–1080.

[66] J. Itten, *Art of color (Kunst der Farbe)*. Otto Maier Verlag, Ravensburg, Germany, 1961.

[67] C. E. Izard, *The Differential Emotions Scale (DES IV-A): A Method of Measuring the Meaning of Subjective Experience of Discrete Emotions*. University of Delaware, 1993.

[68] L. Jacobs, C. Keown, R. Worthley, and K.-I. Ghymn, "Cross-cultural colour comparisons: Global marketers beware!" *International Marketing Review*, vol. 8, no. 3, 1991.

[69] A. Jones and R. Crandall, "Validation of a short index of self-actualization," *Personality and Social Psychology Bulletin*, vol. 12, no. 1, pp. 63–73, 1986.

[70] D. Kahneman, A. B. Krueger, D. A. Schkade, N. Schwarz, and A. A. Stone, "A survey method for characterizing daily life experience: The day reconstruction method," *Science*, vol. 306, no. 5702, pp. 1776–1780, 2004.

[71] C. L. Keyes, "The mental health continuum: From languishing to flourishing in life," *Journal of Health and Social Behavior*, pp. 207–222, 2002.

[72] L. A. King, "The health benefits of writing about life goals," *Personality and Social Psychology Bulletin*, vol. 27, no. 7, pp. 798–807, 2001.

[73] S. Kiritchenko, X. Zhu, and S. M. Mohammad, "Sentiment analysis of short informal texts," *Journal of Artificial Intelligence Research*, vol. 50, pp. 723–762, 2014.

[74] M. Koo, S. B. Algoe, T. D. Wilson, and D. T. Gilbert, "It's a wonderful life: Mentally subtracting positive events improves people's affective states, contrary

to their affective forecasts," *Journal of Personality and Social Psychology*, vol. 95, no. 5, p. 1217, 2008.

[75] E. Kouloumpis, T. Wilson, and J. D. Moore, "Twitter sentiment analysis: The good the bad and the omg!" *International Conference on Weblogs and Social Media*, vol. 11, no. 538-541, p. 164, 2011.

[76] A. D. Kramer, "An unobtrusive behavioral model of gross national happiness," in *Conference on Human Factors in Computing Systems*. ACM, 2010, pp. 287–290.

[77] A. D. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proceedings of the National Academy of Sciences*, p. 201320040, 2014.

[78] J. A. Krosnick, "Response strategies for coping with the cognitive demands of attitude measures in surveys," *Applied Cognitive Psychology*, vol. 5, no. 3, pp. 213–236, 1991.

[79] R. Küller, S. Ballal, T. Laike, B. Mikellides, and G. Tonello, "The impact of light and colour on psychological mood: A cross-cultural study of indoor work environments," *Ergonomics*, vol. 49, no. 14, pp. 1496–1507, 2006.

[80] J. L. Kurtz, "Seeing through new eyes: An experimental investigation of the benefits of photography," *Journal of Basic & Applied Sciences*, vol. 11, p. 354, 2015.

[81] N. D. Lane, M. Mohammod, M. Lin, X. Yang, H. Lu, S. Ali, A. Doryab, E. Berke, T. Choudhury, and A. Campbell, "Bewell: A smartphone application to monitor, model and promote wellbeing," in *International Conference on Pervasive Computing Technologies for Healthcare*, 2011, pp. 23–26.

[82] N. Lathia, "Easy M," https://play.google.com/store/apps/details?id=com.lathia.easym, 2018, accessed: 2018-08-22.

[83] H. Lee, Y. S. Choi, S. Lee, and I. Park, "Towards unobtrusive emotion recognition for affective social communication," in *Consumer Communications and Networking Conference*. IEEE, 2012, pp. 260–264.

[84] J. A. Lee, “SnapAppy android apps on google play,” https://play.google.com/store/apps/details?id=uk.co.jalproductions.snapappy, 2018, accessed: 2018-01-29.

[85] ——, “SnapAppy on the app store,” https://itunes.apple.com/us/app/snapappy/id1203261313, 2018, accessed: 2018-01-29.

[86] J. A. Lee, C. Efstratiou, and L. Bai, “OSN mood tracking: Exploring the use of online social network activity as an indicator of mood changes,” in *International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 2016, pp. 1171–1179.

[87] R. Li Kam Wa, Y. Liu, N. D. Lane, and L. Zhong, “Moodscope: Building a mood sensor from smartphone usage patterns,” in *International Conference on Mobile Systems, Applications, and Services*. ACM, 2013, pp. 389–402.

[88] S. Lichter, K. Haye, and R. Kammann, “Increasing happiness through cognitive retraining,” *New Zealand Psychologist*, vol. 9, no. 2, pp. 57–64, 1980.

[89] H. Lin, W. Tov, and L. Qiu, “Emotional disclosure on social networking sites: The role of network structure and psychological needs,” *Computers in Human Behavior*, vol. 41, pp. 342–350, 2014.

[90] C. L. Lisetti and F. Nasoz, “Using noninvasive wearable computers to recognize human emotions from physiological signals,” *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 1672–1687, 2004.

[91] P. Liu, W. Tov, M. Kosinski, D. J. Stillwell, and L. Qiu, “Do Facebook status updates reflect subjective well-being?” *Cyberpsychology, Behavior, and Social Networking*, vol. 18, no. 7, pp. 373–379, 2015.

[92] LIWC, “LIWC Linguistic Inquiry and Word Count,” http://www.liwc.net/, 2018, accessed: 2018-07-19.

[93] K. Lup, L. Trub, and L. Rosenthal, “Instagram #instasad?: Exploring associations among Instagram use, depressive symptoms, negative social compari-

son, and strangers followed," *Cyberpsychology, Behavior, and Social Networking*, vol. 18, no. 5, pp. 247–252, 2015.

[94] D. D. Luxton, R. A. McCann, N. E. Bush, M. C. Mishkind, and G. M. Reger, "mHealth for mental health: Integrating smartphone technology in behavioral healthcare," *Professional Psychology: Research and Practice*, vol. 42, no. 6, p. 505, 2011.

[95] S. Lyubomirsky and K. Layous, "How do simple positive activities increase well-being?" *Current Directions in Psychological Science*, vol. 22, no. 1, pp. 57–62, 2013.

[96] S. Lyubomirsky and H. S. Lepper, "A measure of subjective happiness: Preliminary reliability and construct validation," *Social Indicators Research*, vol. 46, no. 2, pp. 137–155, 1999.

[97] S. Lyubomirsky, R. Dickerhoof, J. K. Boehm, and K. M. Sheldon, "Becoming happier takes both a will and a proper way: An experimental longitudinal intervention to boost well-being," *Emotion*, vol. 11, no. 2, p. 391, 2011.

[98] Y. Ma, B. Xu, Y. Bai, G. Sun, and R. Zhu, "Daily mood assessment based on mobile phone sensing," in *International Conference on Wearable and Implantable Body Sensor Networks*. IEEE, 2012, pp. 142–147.

[99] C. Macias, T. Panch, Y. M. Hicks, J. S. Scolnick, D. L. Weene, D. Öngür, and B. M. Cohen, "Using smartphone apps to promote psychiatric and physical well-being," *Psychiatric Quarterly*, vol. 86, no. 4, pp. 505–519, 2015.

[100] C. Mackay, T. Cox, G. Burrows, and T. Lazzerini, "An inventory for the measurement of self-reported stress and arousal," *British Journal of Social and Clinical Psychology*, vol. 17, no. 3, pp. 283–284, 1978.

[101] G. MacKerron and S. Mourato, "Happiness is greater in natural environments," *Global Environmental Change*, vol. 23, no. 5, pp. 992–1000, 2013.

[102] A. K. MacLeod, E. Coates, and J. Hetherton, "Increasing well-being through teaching goal-setting and planning skills: Results of a brief intervention," *Journal of Happiness Studies*, vol. 9, no. 2, pp. 185–196, 2008.

[103] T. J. Madden, K. Hewett, and M. S. Roth, "Managing images in different cultures: A cross-national study of color meanings and preferences," *Journal of International Marketing*, vol. 8, no. 4, pp. 90–107, 2000.

[104] A. M. Manago, M. B. Graham, P. M. Greenfield, and G. Salimkhan, "Self-presentation and gender on MySpace," *Journal of Applied Developmental Psychology*, vol. 29, no. 6, pp. 446–458, 2008.

[105] A. Marcengo and A. Rapp, "Visualization of human behavior data: The quantified self," *Big Data: Concepts, Methodologies, Tools, and Applications*, pp. 1582–1612, 2016.

[106] MasaratApp, "Feelic social mood tracker," https://play.google.com/store/apps/details?id=com.masaratapp.feelic, 2018, accessed: 2018-06-11.

[107] C. D. K. Matsumoto and K. Hoashi, "Cyber-social-physical features for mood prediction over online social networks," *Data Engineering and Information Management Forum*, pp. 1–6, 2017.

[108] G. Matthews, D. M. Jones, and A. G. Chamberlain, "Refining the measurement of mood: The UWIST mood adjective checklist," *British Journal of Psychology*, vol. 81, no. 1, pp. 17–42, 1990.

[109] D. McDuff, A. Karlson, A. Kapoor, A. Roseway, and M. Czerwinski, "AffectAura: An intelligent system for emotional memory," in *Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 849–858.

[110] D. McDuff, A. Mahmoud, M. Mavadati, M. Amr, J. Turcot, and R. e. Kaliouby, "AFFDEX SDK: A cross-platform real-time multi-face expression recognition toolkit," in *Conference on Human Factors in Computing Systems: Extended Abstracts*. ACM, 2016, pp. 3723–3726.

[111] D. M. McNair, L. F. Droppleman, and M. Lorr, *Edits manual for the profile of mood states: POMS*. Edits, 1992.

[112] MediaLocal, "Feel Better mood mover," https://feelbetterapp.io/, 2018, accessed: 2018-06-11.

[113] MokriyaLLC, "Moods tracking for better mental health," https://itunes.apple.com/gb/app/moods-tracking-for-better-mental-health/id1023271188, 2017, accessed: 2018-08-13.

[114] J. D. Morris, "Observations: SAM: The Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response," *Journal of Advertising Research*, vol. 35, no. 6, pp. 63–68, 1995.

[115] M. E. Morris, Q. Kathawala, T. K. Leen, E. E. Gorenstein, F. Guilak, M. Labhard, and W. Deleeuw, "Mobile therapy: Case study evaluations of a cell phone application for emotional self-awareness," *Journal of Medical Internet Research*, vol. 12, no. 2, 2010.

[116] W. N. Morris, "A functional analysis of the role of mood in affective systems," *Review of Personality and Social Psychology*, 1992.

[117] A. Muktheeswarar, "Peas mood diary - track emotions," https://play.google.com/store/apps/details?id=com.aumva.peas, 2017, accessed: 2018-06-11.

[118] MyFitnessPal, "MyFitnessPal free calorie counter, diet & exercise journal," https://www.myfitnesspal.com/, 2018, accessed: 2018-08-26.

[119] A. Ortigosa, J. M. Martín, and R. M. Carro, "Sentiment analysis in Facebook and its application to e-learning," *Computers in Human Behavior*, vol. 31, pp. 527–541, 2014.

[120] K. Otake, S. Shimai, J. Tanaka-Matsumi, K. Otsui, and B. L. Fredrickson, "Happy people become happier through kindness: A counting kindnesses intervention," *Journal of Happiness Studies*, vol. 7, no. 3, pp. 361–375, 2006.

[121] E. Owusu, J. Han, S. Das, A. Perrig, and J. Zhang, "ACCessory: Password inference using accelerometers on smartphones," in *Workshop on Mobile Computing Systems & Applications*. ACM, 2012, p. 9.

[122] PacificaLabs, "Pacifica #1 app for anxiety & depression," https://www.thinkpacifica.com/, 2018, accessed: 2018-08-13.

[123] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *International Conference on Language Resources and Evaluation*, vol. 10, no. 2010, 2010, pp. 1320–1326.

[124] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.

[125] A. C. Parks, M. D. Della Porta, R. S. Pierce, R. Zilca, and S. Lyubomirsky, "Pursuing happiness in everyday life: The characteristics and behaviors of online happiness seekers," *Emotion*, vol. 12, no. 6, p. 1222, 2012.

[126] R. L. Payne and C. L. Cooper, *Emotions at work: Theory, research and applications for management*. John Wiley & Sons, 2003.

[127] J. W. Pennebaker, "Conflict and canned meat," *Psychological Inquiry*, vol. 9, no. 3, pp. 219–220, 1998.

[128] ——, "Writing about emotional experiences as a therapeutic process," *Psychological Science*, vol. 8, no. 3, pp. 162–166, 1997.

[129] C. Peterson, "Authentic happiness inventory questionnaire," *Measures Overall Happiness. Michigan: University of Michigan*, 2005.

[130] PewInternetProject, "Social networking fact sheet," http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/, 2014, accessed: 2016-03-31.

[131] R. W. Picard, *Affective computing*. MIT press, 2000.

[132] ——, "Affective computing for HCI," in *Human Computer Interaction*. Citeseer, 1999, pp. 829–833.

[133] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Transactions on Pattern analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, 2001.

[134] M. Pielot, T. Dingler, J. S. Pedro, and N. Oliver, "When attention is not scarce - detecting boredom from mobile phone usage," in *International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 825–836.

[135] L. Piwek and A. Joinson, ""What do they Snapchat about?" Patterns of use in time-limited instant messaging service," *Computers in Human Behavior*, vol. 54, pp. 358–367, 2016.

[136] B. O. Ploog, A. Scharf, D. Nelson, and P. J. Brooks, "Use of computer-assisted technologies (CAT) to enhance social, communicative, and language development in children with autism spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 43, no. 2, pp. 301–322, 2013.

[137] R. Plutchik, "Emotions: A general psychoevolutionary theory," *Approaches to Emotion*, vol. 1984, pp. 197–219, 1984.

[138] J. P. Pollak, P. Adams, and G. Gay, "PAM: A photographic affect meter for frequent, in situ measurement of affect," in *Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 725–734.

[139] D. N. Prata, K. P. Soares, M. A. Silva, D. Q. Trevisan, and P. Letouze, "Social data analysis of Brazilian's mood from Twitter," *International Journal of Social Science and Humanity*, vol. 6, no. 3, p. 179, 2016.

[140] J. Preece and K. Ghozati, "Experiencing empathy online," *The Internet and Health Communication: Experiences and Expectations*, pp. 147–166, 2001.

[141] N. M. Punyanunt-Carter, J. De La Cruz, and J. S. Wrench, "Investigating the relationships among college students' satisfaction, addiction, needs, communication apprehension, motives, and uses & gratifications with Snapchat," *Computers in Human Behavior*, vol. 75, pp. 870–875, 2017.

[142] D. Quercia, L. Capra, and J. Crowcroft, "The social world of Twitter: Topics, geography, and emotions," *International Conference on Weblogs and Social Media*, vol. 12, pp. 298–305, 2012.

[143] D. Quercia, J. Ellis, L. Capra, and J. Crowcroft, "Tracking gross community happiness from tweets," in *Conference on Computer Supported Cooperative Work*. ACM, 2012, pp. 965–968.

[144] D. Quercia, N. K. O'Hare, and H. Cramer, "Aesthetic capital: What makes London look beautiful, quiet, and happy?" in *Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2014, pp. 945–955.

[145] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas, "EmotionSense: A mobile phones based adaptive platform for experimental social psychology research," in *International Conference on Ubiquitous Computing*. ACM, 2010, pp. 281–290.

[146] A. G. Reece and C. M. Danforth, "Instagram photos reveal predictive markers of depression," *EPJ Data Science*, vol. 6, no. 1, p. 15, 2017.

[147] S. L. Rizvi, L. A. Dimeff, J. Skutch, D. Carroll, and M. M. Linehan, "A pilot study of the DBT coach: An interactive mobile phone application for individuals with borderline personality disorder and substance use disorder," *Behavior Therapy*, vol. 42, no. 4, pp. 589–600, 2011.

[148] M. Roshanaei, R. Han, and S. Mishra, "Features for mood prediction in social media," in *International Conference on Advances in Social Networks Analysis and Mining*. IEEE, 2015, pp. 1580–1581.

[149] J. A. Russell, "Affective space is bipolar," *Journal of Personality and Social Psychology*, vol. 37, no. 3, p. 345, 1979.

[150] ——, "Core affect and the psychological construction of emotion," *Psychological Review*, vol. 110, no. 1, p. 145, 2003.

[151] ——, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, p. 1161, 1980.

[152] J. A. Russell, A. Weiss, and G. A. Mendelsohn, "Affect grid: A single-item scale of pleasure and arousal," *Journal of Personality and Social Psychology*, vol. 57, no. 3, p. 493, 1989.

[153] D. Ruths and J. Pfeffer, "Social media for large studies of behavior," *Science*, vol. 346, no. 6213, pp. 1063–1064, 2014.

[154] R. M. Ryan and C. Frederick, "On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being," *Journal of Personality*, vol. 65, no. 3, pp. 529–565, 1997.

[155] C. D. Ryff and C. L. M. Keyes, "The structure of psychological well-being revisited," *Journal of Personality and Social Psychology*, vol. 69, no. 4, p. 719, 1995.

[156] H. Sakawa, F. Ohtake, and Y. Tsutsui, "Activity, time, and subjective happiness: An analysis based on an hourly web survey," *The Institute of Social and Economic Research Discussion*, 2015.

[157] P. Sanches, A. Janson, P. Karpashevich, C. Nadal, C. Qu, C. Dauden Roquet, M. Umair, C. Windlin, G. Doherty, K. Höök, and C. Sas, "HCI and affective health: Taking stock of a decade of studies and charting future research directions," *Conference on Human Factors in Computing Systems*, 2019.

[158] G. M. Sandstrom, N. Lathia, C. Mascolo, and P. J. Rentfrow, "Putting mood in context: Using smartphones to examine how people feel in different locations," *Journal of Research in Personality*, vol. 69, pp. 96–101, 2017.

[159] A. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," in *Conference on Affective Computing and Intelligent Interaction*. IEEE, 2013, pp. 671–676.

[160] C. Sas, A. Dix, J. Hart, and R. Su, "Dramaturgical capitalization of positive emotions: The answer for Facebook success?" in *Conference on People and*

Computers: Celebrating People and Technology. British Computer Society, 2009, pp. 120–129.

[161] C. Sas, T. Fratczak, M. Rees, H. Gellersen, V. Kalnikaite, A. Coman, and K. Höök, "AffectCam: Arousal-augmented sensecam for richer recall of episodic memories," in *Conference on Human Factors in Computing Systems: Extended Abstracts*. ACM, 2013, pp. 1041–1046.

[162] K. R. Scherer, "What are emotions? And how can they be measured?" *Social Science Information*, vol. 44, no. 4, pp. 695–729, 2005.

[163] M. E. Seligman and M. Csikszentmihalyi, *Positive psychology: An introduction*. American Psychological Association, 2000, vol. 55, no. 1.

[164] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 410–417, 2010.

[165] K. M. Sheldon, T. Kasser, K. Smith, and T. Share, "Personal goals and psychological growth: Testing an intervention to enhance goal attainment and personality integration," *Journal of Personality*, vol. 70, no. 1, pp. 5–31, 2002.

[166] L. Shen, M. Wang, and R. Shen, "Affective e-learning: Using "emotional" data to improve learning in pervasive learning environment," *Journal of Educational Technology & Society*, vol. 12, no. 2, 2009.

[167] S. Shiffman, A. A. Stone, and M. R. Hufford, "Ecological momentary assessment," *Annual Review of Clinical Psychology*, vol. 4, pp. 1–32, 2008.

[168] P. E. Sifneos, "The prevalence of 'alexithymic' characteristics in psychosomatic patients," *Psychotherapy and Psychosomatics*, vol. 22, no. 2-6, pp. 255–262, 1973.

[169] N. L. Sin and S. Lyubomirsky, "Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis," *Journal of Clinical Psychology*, vol. 65, no. 5, pp. 467–487, 2009.

[170] A. Sonderegger, K. Heyden, A. Chavaillaz, and J. Sauer, "AniSAM & AniAvatar: Animated visualizations of affective states," in *Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 4828–4837.

[171] O. Sourina and Y. Liu, "A fractal-based algorithm of emotion recognition from EEG using arousal-valence model," in *Biosignals*, 2011, pp. 209–214.

[172] C. D. Spielberger, G. Jacobs, S. Russell, and R. S. Crane, "Assessment of anger: The state-trait anger scale," *Advances in Personality Assessment*, vol. 2, pp. 159–187, 1983.

[173] B. Spring, M. Gotsis, A. Paiva, and D. Spruijt-Metz, "Healthy apps: Mobile devices for continuous monitoring and intervention," *IEEE Pulse*, vol. 4, no. 6, pp. 34–40, 2013.

[174] Stephanie, "Z-Score definition, formula and calculation," http://www.statisticshowto.com/probability-and-statistics/z-score/, 2018, accessed: 2018-06-13.

[175] J. Sykes and S. Brown, "Affective gaming: Measuring emotion through the gamepad," in *Conference on Human Factors in Computing Systems: Extended Abstracts*. ACM, 2003, pp. 732–733.

[176] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, "User-level sentiment analysis incorporating social networks," in *International Conference on Knowledge Discovery and Data Mining*. ACM, 2011, pp. 1397–1405.

[177] J. P. Tangney, *The test of self-conscious affect*. George Mason Univ., 1989.

[178] S. E. Taylor, "Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis," *Psychological Bulletin*, vol. 110, no. 1, p. 67, 1991.

[179] R. E. Thayer, "Factor analytic and reliability studies on the activation-deactivation adjective check list," *Psychological Reports*, vol. 42, no. 3, pp. 747–756, 1978.

[180] P. Thiffault and J. Bergeron, "Monotony of road environment and driver fatigue: A simulator study," *Accident Analysis & Prevention*, vol. 35, no. 3, pp. 381–391, 2003.

[181] E. R. Thompson, "Development and validation of an internationally reliable short-form of the positive and negative affect schedule (panas)," *Journal of Cross-Cultural Psychology*, vol. 38, no. 2, pp. 227–242, 2007.

[182] ThrivePort, "MoodKit CBT app," http://www.thriveport.com/products/moodkit/, 2018, accessed: 2018-08-13.

[183] ——, "MoodNotes thought journal, mood diary, CBT app," http://moodnotes.thriveport.com/, 2015, accessed: 2018-06-11.

[184] Timehop, "Timehop," https://www.timehop.com/, 2018, accessed: 2018-09-04.

[185] O. Tsur, D. Davidov, and A. Rappoport, "A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews," in *International Conference on Weblogs and Social Media*, 2010, pp. 162–169.

[186] TwoAppStudio, "MoodCast smart mood diary for Android," https://2appstudio.com/moodcast/, 2017, accessed: 2018-06-11.

[187] S. Utz, N. Muscanell, and C. Khalid, "Snapchat elicits more jealousy than Facebook: A comparison of Snapchat and Facebook use," *Cyberpsychology, Behavior, and Social Networking*, vol. 18, no. 3, pp. 141–146, 2015.

[188] T. Vogel, "Year in Pixels mood and emotion tracker," https://play.google.com/store/apps/details?id=ar.teovogel.yip, 2018, accessed: 2018-06-11.

[189] T. Vos *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A systematic analysis for the global burden of disease study 2013," *The Lancet*, vol. 386, pp. 743–800, 2015.

[190] W. R. Walker, J. J. Skowronski, and C. P. Thompson, "Life is pleasant–and memory helps to keep it that way!" *Review of General Psychology*, vol. 7, no. 2, p. 203, 2003.

[191] N. Wang, M. Kosinski, D. Stillwell, and J. Rust, "Can well-being be measured using Facebook status updates? Validation of Facebook's Gross National Happiness Index," *Social Indicators Research*, vol. 115, no. 1, pp. 483–491, 2014.

[192] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "StudentLife: Using smartphones to assess mental health and academic performance of college students," in *Mobile Health*. Springer, 2017, pp. 7–33.

[193] D. Watson and L. A. Clark, "The PANAS-X: Manual for the positive and negative affect schedule-expanded form," *Department of Psychological & Brain Sciences Publications*, 1999.

[194] D. Watson and A. Tellegen, "Toward a consensual structure of mood," *Psychological Bulletin*, vol. 98, no. 2, p. 219, 1985.

[195] D. Watson, L. A. Clark, and G. Carey, "Positive and negative affectivity and their relation to anxiety and depressive disorders," *Journal of Abnormal Psychology*, vol. 97, no. 3, p. 346, 1988.

[196] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The PANAS scales," *Journal of Personality and Social Psychology*, vol. 54, no. 6, p. 1063, 1988.

[197] F. Weber, "Daylio mood tracker and micro diary," https://daylio.webflow.io/, 2018, accessed: 2018-06-11.

[198] J. Weiser, *Phototherapy techniques: Exploring the secrets of personal snapshots and family albums*. PhotoTherapy Centre Vancouver, BC, 1999.

[199] G. N. Yannakakis and A. Paiva, "Emotion in games," *Handbook on Affective Computing*, pp. 459–471, 2014.

[200] A. Zenonos, A. Khan, G. Kalogridis, S. Vatsikas, T. Lewis, and M. Sooriyabandara, "HealthyOffice: Mood recognition at work using smartphones and wearable sensors," in *International Conference on Pervasive Computing and Communication Workshops*. IEEE, 2016, pp. 1–6.

[201] ZNStudio, "Emotion Gram mood tracker," https://play.google.com/store/apps/details?id=diary.daylio.mood.tracker, 2016, accessed: 2018-06-11.