Does Smartphone Use Drive our Emotions or vice versa? A Causal Analysis

Zhanna Sarsenbayeva[1], Gabriele Marini[1], Niels van Berkel[2], Chu Luo[1], Weiwei Jiang[1], Kangning Yang[1], Greg Wadley[1], Tilman Dingler[1], Vassilis Kostakos[1], Jorge Goncalves[1]
[1] The University of Melbourne, Melbourne, Australia, firstname.lastname@unimelb.edu.au
[2] Aalborg University, Aalborg, Denmark, nielsvanberkel@cs.aau.dk

ABSTRACT
In this paper, we demonstrate the existence of a bidirectional causal relationship between smartphone application use and user emotions. In a two-week long in-the-wild study with 30 participants we captured 502,851 instances of smartphone application use in tandem with corresponding emotional data from facial expressions. Our analysis shows that while in most cases application use drives user emotions, multiple application categories exist for which the causal effect is in the opposite direction. Our findings shed light on the relationship between smartphone use and emotional states. We furthermore discuss the opportunities for research and practice that arise from our findings and their potential to support emotional well-being.

Author Keywords
Smartphones, emotions, mobile interaction, mobile application use, emotion detection, emotional well-being

INTRODUCTION
Emotions are integral to human behaviour and influence the way we think and act [53]. They incline people to respond effectively to environmental challenges and opportunities [31]. Thus, emotions are inextricably intertwined with people's interactions with the world. This is a *bidirectional* influence: our experiences in the world engender emotional reactions, and in turn these emotions shape our behaviour and interactions [78]. In this paper, we investigate the bidirectional causal relationship between emotions and the most ubiquitous interactive device: smartphones.

As digital technology use becomes an increasingly important component of our daily experience, it is to be expected that the experience of using technology will also sometimes be imbued with emotion. Indeed, prior work has long shown this to be the case, with the earliest results in HCI research emphasising the importance of emotion in user experience [58] and the adoption of technology [43], and asserting the importance of designing technology to create desirable emotional experiences [67]. Recently, researchers have studied this in more specific contexts. For example, the use of digital technologies as diverse as smartphones [61], videogames [20], social media [48], online videos [64], music streaming services [49] and virtual reality [94] are known to influence emotions.

Recent research has also examined the 'inverse' causal direction: how emotional states influence technology use. States of joy and anger have been shown to affect interaction, such as increase joystick movements, speed, and acceleration compared to states of sadness [4]. Even typing speed is affected by our emotional state: Khanna and Sasikumar [47] found that when experiencing negative emotional states users type slower compared to when experiencing positive states. Likewise, a user's emotional states influence interaction with mobile devices [13]. For example, Cairns *et al.* show in their study that users in a negative emotional state are more likely to make errors when completing a simple task of number entry on a tablet [13]. Moreover, Mottelson and Hornbæk [63] demonstrate that a user's emotional state has a behavioural effect on mobile interaction. The authors show that the participants with neutral affective state had a higher acceleration when performing touch tasks as compared to participants with positive affective state.

Given the prominence of smartphones, and the evidence on the effect of emotional state on smartphone use, we set out to further investigate the relationship between smartphone use and users' emotions. In particular, we address the challenge of understanding whether specific emotions drive particular smartphone use, and/or whether smartphone use leads to certain emotions. A study by Mehrotra *et al.* [61] was one of the first to identify a causal impact of users' emotions on mobile interaction. The study collected participants' self-reported emotional states and mapped these to their phone use behaviour in the subsequent hour. We extend this work by 1) utilising a non-administered *in situ* method for the passive and continuous assessment of emotions from facial expressions during application use, and 2) applying the Convergent Cross Mapping (CCM) [89] method, published in *Science* in 2012 and recently extended to the HCI domain to combine, visualise, and evaluate multiple participants [95], to carefully investigate the *bidirectional* causal relationship between user emotions and smartphone use. This is made possible due to

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376163

the fine granularity of our data, unlike self-report data which is often sparse.

To conduct the study, we developed ‘ApplicationTracker’ — an Android application used to collect information on the user’s application use as well as their emotional state, as derived from the user’s facial expressions captured with the smartphone’s front-facing camera. ‘ApplicationTracker’ utilises the AWARE Framework [28] to collect smartphone application use and the Affectiva Emotion SDK [1] for detecting emotions from facial expressions. This combination allows us to record and analyse emotional states during smartphone interaction. Our analysis shows that, overall, phone use drives certain emotions rather than the other way around. Furthermore, we identify specific application categories which actually drive users’ emotions.

The contribution of our work is three-fold:

1. *Methodologically*, we use simultaneous and continuous mapping of users’ facial expressions and application use traces to study the relationship between emotional state and smartphone application use.

2. *Analytically*, we show that smartphone application use drives user emotions more often than the other way around, and we shed more light on this relationship.

3. *Conceptually*, we highlight the implications of our findings for both research and practice. In particular, we argue that understanding the relationship between application use and emotions has the potential to promote and facilitate emotional well-being through emotion regulation practices [40].

RELATED WORK

Emotions and Their Effects on Human Behaviour

Emotions are part of everyday experience and are causally intertwined with behaviour [53]. For example, prior literature has established a positive effect of enjoyment on performance in education [52]. In particular, Larson *et al.* [52] found that students who enjoyed their projects, received a higher grade upon finishing the project ($r = 0.44$). Furthermore, positive emotions have been shown not only to improve performance, but also to facilitate creative, flexible, and divergent thinking and hence, efficient-problem solving [76]. Positive emotions have also been shown to increase focus attention on the task [70]. Negative emotions (*e.g.,* sadness), in contrast to positive emotions, have been shown to negatively influence human ability to concentrate [23]. Moreover, medical students who were induced with negative emotions took significantly longer time to diagnose patients based on an x-ray as compared to students who were induced with positive emotions [45].

Given the established effect of emotions on human cognition and behaviour, some researchers have investigated the effect of emotions on mobile interaction. For example, Mottelson *et al.* found that participants exposed to positive emotional stimuli were more precise in tapping tasks as compared to participants with neutral-induced affective state [63]. In addition, Coombes *et al.* [21] have investigated the effect of neutral, unpleasant, and pleasant pictures on a square-tracing smartphone task. The authors found that the participants were quicker to perform the task but made more errors after viewing unpleasant pictures compared to the other two conditions [21].

Emotion Detection Mechanisms

Emotion detection has recently gained significant attention across different scientific disciplines, including Computer Science [107]. For example, Epp *et al.* suggested using typing patterns on a standard computer keyboard to detect user emotions [26]. The authors modelled six emotions (confidence, hesitance, nervousness, relaxation, sadness, and tiredness) with accuracy rates ranging between 77.4% and 87.8%. Savva and Bianchi-Berthouze [84] propose a system to recognise user emotions while playing a Wii tennis game.

Another method of measuring emotions is through self-reports. A common way to assess a person’s emotional state is to ask them to rate their feelings according to the Valence-Arousal dimensions [72]. Techniques such as the self-assessment mannequin (SAM) [50], circumplex model of affect [79], photographic affect meter (PAM) [74], and positive and negative affect schedules (PANAS) [101] are widely used to collected self-reported emotional state.

Finally, previous work has explored the use of physiological attributes to detect emotions. Examples include the use of facial expressions [18], skin conductance [72], and heart rate variability [55]. Facial expressions not only offer insights into a person’s emotional state, but they can also provide additional physiological data (*e.g.,* heart rate, heart rate variability, skin colour changes, and respiration rate) which can be used to detect emotions [18]. For example, Poh *et al.* used a webcam to extract various features (heart rate, respiratory rate, and heart rate variability) to detect emotions from the captured facial images [73]. Although some scientists argue that facial expressions do not reflect the people’s true emotional feelings [29], several researchers agree that facial expressions can serve as indicators of people’s mental state [30, 57] and reflect emotions as a normal physiological response [106, 25]. For example, a person’s visible smile is typically an indicator of happiness, appreciation, and desire [106, 80].

Detecting Emotions with Smartphones

Previous work has shown that smartphone interaction behaviour and sensor data can be used to detect emotional state of the user [12, 54]. For instance, a number of studies have assessed mental health using smartphone sensor data by looking at phone use and contextual data [12, 56], facial expressions [65], social interaction [54], speech patterns [17, 19], as well as sleep and physical activity [77, 100].

Gao *et al.* [32] suggested using mobile touch to detect user’s emotional state. The authors extracted participants fingerstroke features while playing *Fruit Ninja*. The researchers achieved an 89.7% accuracy from touch strokes as compared to self-assessed emotional states. However, it is difficult to generalise these results due to the specificity of the task and the limited sample size (15 participants). LiKamWa *et al.* presented MoodScope – an application which detect user’s mood [56]. In this study, the authors collected contextual data

[1] https://knowledge.affectiva.com/

(phone calls, text messages, emails, application use, web browsing histories, and location data) as well as self-reported mood states [56]. MoodScope achieved 66% accuracy for predicting participants' daily average mood with phone calls and application categories being the most useful features in prediction models [56]. Similarly, Sun *et al.* utilised sensor data, event data, content data, and application use data to predict user's emotions [90]. Zhang *et al.* employed application use data, contextual information, and sensor data to detect user's compound emotions [107]. Hung *et al.* [44] used call logs and application use together with user's self-reported emotions to predict negative emotion.

These examples show different efforts in the research community to use an off-the-shelf smartphone to sense emotional state of the user. In this study, we collect *in situ* emotions derived from facial expressions of our participants during their smartphone application use using a validated tool [87]. This passive sensing approach provides us with the insights of people's instant emotional reactions to applications in a real world scenario with a high degree of granularity without being burdensome to the participants.

Application Use Behaviour

Studies on application use behaviour have gained popularity in recent years [62, 60]. Application use patterns have been used to understand people's behaviour [71]. For example, Silva *et al.* [85] show cultural differences in eating and drinking habits from *FourSquare* use patterns. The authors demonstrate that in most of the cases, cities from the same country tend to have similar drinking and fast food habits [85]. Goodrich and De Mooij [36] investigate cultural influences of social media on consumerism. They also found that human-sources of information is more valuable in cultures of short-term orientation; whereas in cultures of long-term orientation, fact-based information (*e.g.,* search-engines) are more important [36]. Song *et al.* [86] argue that application use information can provide opportunities to optimise smartphone systems. The authors present a personalised optimisation framework that enables prediction of future app launch; hence, reducing unnecessary app restarts by 78.4%.

There have also been several attempts to derive users' emotional state from application use data. In a study by Visuri *et al.* [99], the authors map users' affective state to applications use. They found that when experiencing positive affect, the users tend to use Media, Games, Maps and Travel applications. However, when the affective state is negative, the users have a tendency to use Internet browsers or Social media apps [99]. A study by Mehrotra *et al.* [61] is most closely related to our study. The authors examined the causal relationship between the users' emotional state and smartphone interaction. They collected smartphone use data (*e.g.,* app use, micro-interactions with the device, communication patterns, and notifications) and used the Experience Sampling Method to collect users' self-reported emotional states. Their findings suggest that user's activeness levels have a positive impact on the use of music applications, and that an increase in stress levels significantly reduces the use of communication applications.

In this paper, we extend the work by Mehrotra *et al.* [61], by investigating the bidirectional causal relationship between users' emotional state and their smartphone application use. However, we utilise emotions derived from facial expressions (not using self-report data collection methods due to their sparsity) and investigate causal relationship between the instant effect of the application use on emotional state of the user through Convergent Cross Mapping (CCM) analysis.

METHOD

Software – 'ApplicationTracker'

We developed a standalone Android application called 'ApplicationTracker' that uses the AWARE framework [28] to collect smartphone use data. In particular, the application logs unlock events and the foreground applications used during the session following the unlock event.

We integrated the Affectiva Android API [59] into 'ApplicationTracker' to track user's emotional expressions with the device's front-facing camera during smartphone use. Thus, the software continuously logs confidence values (between 0 to 100) for the emotions anger, contempt, disgust, fear, joy, sadness, and surprise, estimated from the user's facial expressions at the default 30 frames per second. The software starts logging user emotions upon unlock and stops logging the emotions when the screen is locked. All emotion and application use data is recorded with a corresponding timestamp. The software runs as a service listening for unlock events and has a simple interface consisting of a button that can be used to start the service upon deployment.

We also collected self-reported emotional states of our participants (valence and arousal) through Experience Sampling Method (ESM) questionnaires [22] in order to triangulate against and check the accuracy of the Affectiva data. The questionnaires were administered 6 times per day between 9:00 am and 9:00 pm at random times with a minimal interval of 1 hour between two consecutive questionnaires [96].

Affectiva Emotion-Detection API Validation

We evaluated Affectiva's performance on emotion detection on three publicly available databases in order to increase our confidence on the reliability of the API in detecting the different emotions. The human models posing for all these databases were trained by Facial Action Coding System (FACS) [24] experts to express basic emotions. The three datasets used were:

- The Amsterdam Dynamic Facial Expression Set (ADFES) [97] – Database containing both dynamic (videos) and static (pictures) of facial expressions of anger, contempt, disgust, fear, happiness, sadness, and surprise from 22 people. The pictures were validated by 119 non-experts judges.
- The Warsaw Set of Emotional Facial Expression Pictures (WSEFEP) [68] – Database consisting of 210 images of 30 people, expressing anger, disgust, fear, happiness, sadness, and surprise. The pictures were validated by a large number of non-expert judges ($N = 1362$).

- The Radboud Faces Database (RaFD) [51] – Database that contains 536 pictures of faces expressing basic emotions (anger, contempt, disgust, fear, happiness, sadness, and surprise) from 67 people. The pictures were validated by 238 non-expert judges.

We used a custom automated way to feed the images from the databases to the Affectiva API and recorded the resulting emotion classification. We initially analysed each emotion individually to check Affectiva's detection accuracy. We found that Affectiva was more accurate at detecting certain emotions than others, *e.g.,* surprise (accuracy of 97.46%), joy (accuracy of 94.17%), disgust (accuracy of 89.92%), and contempt (accuracy of 81.36%). Whereas, Affectiva was relatively inaccurate in detecting anger (50.82%) and fear (7.5%). Hence, we decided to remove these two emotion classifications from our user-study dataset to increase the reliability of our findings. After the removal of these two emotions, our results show that Affectiva accurately detected emotions with an overall accuracy of 85%.

Participants and Procedure

We recruited 30 participants (15 Female, 15 Male) aged between 20 and 45 (M = 29, SD = 6.07) via our university's mailing lists and through snowball recruiting. Participants were required to own an Android-based smartphone to enrol in the experiment. In addition, our participants had a diverse range of educational background (*e.g.,* Accounting, Biomedicine, Computer Science, Education, Linguistics, Neuroscience) and each participant was assigned a unique anonymous ID. Moreover, almost all participants (29 out of 30) had a standard 5 days work week with 2 days of weekend.

We held individual intake sessions for each participant. Upon arrival to our lab, we briefed them about the purpose of the experiment, and received their consent agreeing to participate in our study. We then installed the software on their smartphone and explained how it worked. The data collection phase lasted for two weeks, during which participants were instructed to use their phones as usual. App usage and facial expressions data were stored locally on the user's smartphone for privacy reasons.

After the data collection was finished, we invited participants for individual debriefing sessions. During these sessions we retrieved the data from each participant's device, uninstalled the software, and held semi-structured exit interviews to inquire about their perceptions of their emotions and applications use. Every participant was rewarded with a $10 gift voucher for their participation. The experimental design was approved by the Ethics Committee of our university.

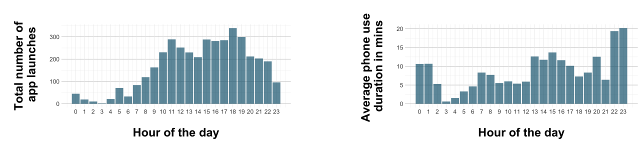
Software Power Consumption Test

We conducted a test on a smartphone to quantify the power consumption of our software. We selected a Samsung S7 with Android 7.0 as the experiment device (battery capacity $3000mAh$). We installed our software and launched 248 unlock events leaving the screen on for 30 seconds each time. This number of unlock events is substantially higher than the daily use of a regular user ($10-200$ unlocks per day) [27]. To provide a baseline, we also recorded the energy consumption of the same events when the software was not running on the device. For the 248 unlock events solely, the devices used 8% of its battery capacity ($240mAh$). To execute our software after 248 unlocks, the device uses an additional 11% ($330mAh$) of its battery capacity. Considering the large number of software executions, the energy consumption of our software in practical use is minor (approximately 1% power per 23 unlocks) given that the users unlock their phones $10-200$ times on average per day [27].

RESULTS

General Descriptive Statistics

Overall, we collected 502,851 data points where application use was matched to the corresponding emotion records. We analysed our data to count the total number of application *launches* per hour-of-day and visualised it in Figure 1-(a). The figure shows that the number of application launches grows towards the evening. These results are in line with previous findings [10]. We then visualised average phone use *duration* per hour in Figure 1-(b). The figure shows that the highest average duration of phone use occurs between 22:00 and midnight while the lowest phone use occurs between 2:00 am and 7:00 am.



(a) (b)

Figure 1: Statistics on phone usage

We then analyse the fluctuations of our participants' emotion throughout the week. As joy and sadness have been shown to follow day-of-week pattern [88, 35], we chose these emotions to demonstrate that the trend also holds in our study. The fluctuations of emotions, (a) joy and (b) sadness, are presented in Figure 2. For the data in Figure 2, we averaged and smoothed the data from all our participants and grouped per day of the week and find a pattern as shown in previous work [88]. The apparent trend in joy and sadness over the course of the week provides confidence in the detection features of the Affectiva API. We can see that towards the weekend participants' joy increases, while on weekdays it drops; whereas sadness drops towards the weekend and is at its peak on weekdays. Our results are in line with previous research [88], and provide additional confidence in the accuracy of the emotion measurements collected via the Affectiva API.

We also collected participants' self-reported affective states through ESM questionnaires. We followed a protocol presented by Van Berkel *et al.* [96] and constructed the ESM questionnaire according to Russel's Circumplex Model [79]. The ESM questionnaire contained two Likert-scale questions: *Miserable–Pleased* and *Sleepy–Aroused*. The use of this scale allowed participants to indicate the intensity of their current emotion rather than just a label. This aligns more closely with confidence values as obtained through Affectiva and allows for the subsequent CCM analysis.

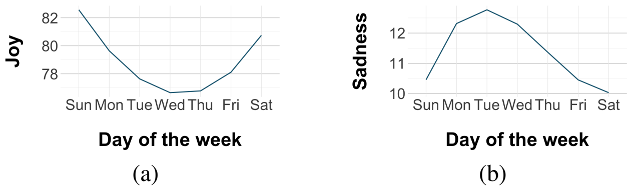


Figure 2: Daily fluctuations for joy and sadness throughout the week (smoothed conditional means)

This enabled us to provide additional validation of the Affectiva data beyond the tests described above. Note that this validation is limited since the ESM data is explicitly entered by participants only up to 6 times per day. Because the frequencies of the two datasets are substantially different, and they also capture slightly different behavioural aspects, our analysis proceeds as follows. We consider the valence data from the ESMs, and we calculate the mean valence over the preceding minute as generated by the Affectiva. Next, we consider time segments where both ESM and Affectiva data were recorded, and we calculate the correlation of ESM with Affectiva data. We find that ESM valence positively correlates with the mean Affectiva valence values ($r = 0.3, p < 0.01$). This relationship varies in proportion to the time window we consider. For example, when we consider a 1-hour window, the correlation drops to 0.11. When we consider a 5-minute window, the correlation goes up to 0.23, and for 1-minute window it is 0.3. The results validate our expectations since we expect a stronger correlation as the time window closes in on the ESM response time, and weaker correlation as we consider an increasing time period around the ESM instance.

Convergent Cross Mapping

Next, we applied the Convergent Cross Mapping (CCM) method introduced by Sugihara *et al.* [89] to further analyse our data. CCM is a method commonly used to distinguish causality from correlation developed for time series variables. CCM is a newly developed alternative to Granger causality (GC), which is a traditional approach to investigate causal relationship between two variables in a time series. GC has been used empirically to find patterns of correlations in cases where the system is purely stochastic [37]. To clarify, CCM is not a version of GC, but rather both CCM and GC are methods for identifying cause and effect in timeseries. Whereas GC is used for the analysis of clearly separable variables, CCM considers timeseries data from a complex and dynamic systems perspective, and investigates the relationships between variables in a system that is not entirely random. Unlike GC, CCM uses the notion of convergence to distinguish causation from correlation.

Interaction between user and technology can be influenced by many different factors such as environmental and social context, mood, friends, weather, and applications use which potentially can change by themselves and in relation to each other in nonlinear ways. For example, when people interact with their smartphones, their application use might lead them to feel certain emotions; and vice versa when feeling certain emotions, people might use specific applications on their phone. Such a behavioural loop can be investigated from the perspective of coupled dynamic systems in order to understand the causality and its direction between emotions and application use. These insights can be used in the design of *e.g.,* digital emotional regulation applications and to increase our understanding of the effect of digital devices on end users. Highly established work in the HCI field, such as Don Norman's 'Gulf of execution and evaluation', shows that human behaviour can be described as a complex system [16, 33, 66].

As we seek to identify causal relationships in our time-series based data, CCM presents itself as an appropriate and robust method which avoids some of the restrictive assumptions made by GC. We used the "rEDM" R package by Ye *et al.* [105] to conduct our analysis and follow the analysis approach presented in [95].

We analysed the causality between emotions and phone use in terms of 1) the number of applications launched and 2) app usage duration. We summarise our results in Tables 1 and 2 as well as in Figures presented in the Supplementary Materials. Tables 1 and 2 are colour-coded according to the direction of causal relationship and its effect size: darker colours indicate a stronger effect; blue indicates that phone use drives emotions, and orange indicates that emotions drive phone use.

The procedure we follow is identical for all tests we report here: we group the emotion data and app usage data into hourly time-series, and each time-series is fed into the CCM algorithm. Specifically, Affectiva generates one record every 250 milliseconds, containing the estimated value for each of the 7 emotions that it tracks. Therefore, over a period of time, we collect an equal number of readings for each of these emotions. We average these values, per participant per emotion, into 1-hour periods, and also calculate for each period how many applications were launched, and how much time was spent using these applications. Therefore, for each 1-hour period we have an estimated value for 'mean joy', 'mean sadness', etc., as well as the total number of application launches and total duration. We did not apply any first-differencing to the data as CCM has been applied to records in which samples are not evenly spaced in time (*e.g.,* Van Nes *et al.* [98]), but the state-space cross-mapping concept demands that the variables being compared are sampled at contemporaneous points. CCM does not assume stationarity, hence there is no need to detrend the observed time series [41]. These values form our time series that we feed into the CCM analysis.

CCM Parameters

We grouped the emotion data and application usage data into hourly time-series, hence used the one hour time delay as it has been shown that emotions persist over time [6, 7]. Using shorter time-slots (*e.g.,* 1-minute or 5-minute) results in most slots having 0 values, and therefore the data becomes less meaningful. We also argue that hourly slots do not pose a challenge in terms of detecting transient emotions, because CCM actually identifies cause-and-effect by considering all values in the time series in varying window sizes. In addition, emotions in each 1-hour segment do not "cancel each other out" (*e.g.,* joy and sadness are aggregated separately over the time period) as both our data collection (Affectiva) and data analysis consider each emotion independently. Essentially, having "richer" observation periods allows for more confidence in the results. Finally, we confirm that the results from our analysis were matched against interview material with our participants, and therefore our findings are triangulated. Nevertheless, we agree that future work should consider analysing such data at different time lag configurations as it has been shown to improve the results of the CCM analysis [104].

To identify the optimum value for E (Embedding Dimension) we used simplex projection as was recommended in [104]. The idea is to use a set of E lagged values of a variable in order to remodel the behaviour of a dynamic system in E-space (this is consistent with Takens theorem [91]). Each point in E-space is formed using the variables' E lags and these points construct an 'attractor' or an 'attractor manifold' that define the system's evolution. Then, for each point we find the E+1 nearest neighbours and project them to forecast future values in order to evaluate the quality of reconstruction. This forecasting power is calculated as the correlation between the observed and predicted values (rho value). The optimum E value is the one that maximises this correlation as it provides the best out-of-sample predictions of the future.

Finally, the convergence is visually verified by ensuring that as the number of points on the manifold increases, the strength of the causal effect increases and then plateaus. The convergence is visualised in the left frame of Figure 3. Then, for each participant we visualise the outcome of the convergence and position it on the x-y axes. This final graph provides an overview of the causality analysis for the two respective variables.

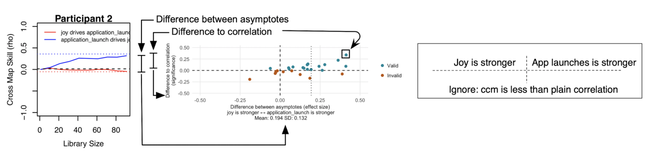


Figure 3: A visualisation of the difference between asymptotes and the difference to correlation. These two values represent x and y coordinates of a datapoint in the summary scatterplot.

Application Launches vs. User Emotions

We investigate the bidirectional causal relationship between the number of app *launches* and emotional state of the user to see if applications launches lead participants to experience certain emotions, and/or vice-versa.

With CCM, the direction of causality is established by looking at how well variable 1 can forecast variable 2, and comparing it to how well variable 2 can forecast variable 1. This comparison establishes the overarching direction of causality between two variables. Subsequently, the strength of this forecasting ability is compared to the raw correlation between the two variables. If the forecasting ability is lower than the plain correlation, then the result is discarded. If the forecasting ability is greater than the plain correlation, then the result is retained. Then, we calculate the overall direction of causality by averaging the valid data points (*i.e.,* each individual participant).

Metric	Contempt	Disgust	Joy	Sadness	Surprise
Total Apps Launch	0.203 (0.140)	0.174 (0.205)	0.194 (0.132)	0.108 (0.214)	0.202 (0.311)
Communication Apps Launch	0.226 (0.261)	0.059 (0.414)	0.147 (0.224)	-0.002 (0.391)	0.260 (0.383)
Social Apps Launch	0.305 (0.403)	0.092 (0.429)	0.479 (0.153)	-0.083 (0.555)	-0.015 (0.288)
Work Apps Launch	0.484 (0.185)	0.251 (0.129)	0.257 (0.252)	0.070 (0.555)	0.088 (0.589)
Entertainment Apps Launch	0.202 (0.348)	0.239 (0.542)	0.487 (0.481)	0.262 (0.847)	-0.151 (0.621)

Table 1: Effect Sizes (and SD) for Causality between Application Launch and Emotions. Blue: phone use drives emotions; Orange: emotion drives phone use.

We visualise the results in Figure 5 (Supplementary Materials); however a part of the graph is presented in Figure 4. Each participant in the graph is represented as a dot. For each participant, the CCM algorithm determines whether variable 1 (*e.g.,* joy) is driven by a variable 2 (*e.g.,* the number of app launches), or vice versa. The direction and magnitude of this effect for each participant is indicated by the position of the data point on the x-axis. If the point is to the left of the y-axis, it suggests that the effect is in a certain direction (joy drives application launches); if the point is to the right, then the effect is in the opposite direction (application launches drive joy). The distance from the x-axis indicates how much stronger this forecasting ability is than plain correlation. Finally, the dashed vertical line indicates the mean effect size of causality across all participants. When interpreting the data points, we ignore the red dots as they are labelled by CCM as invalid: for these dots, CCM is not able to provide explanatory power that is greater than a plain correlation.

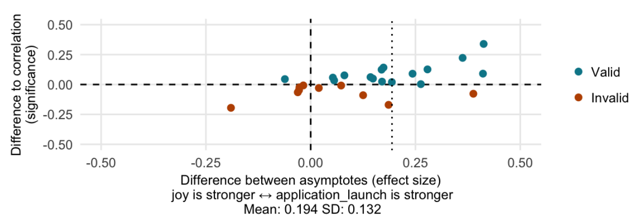


Figure 4: CCM graph visualising the causal relationship between application launches and joy)

In the case of Figure 5 (Supplementary Materials), the prevalence of points to the right of the graphs' origin indicates that application launch drives emotions for the majority of the participants. We note that "drive" is a concept from complex systems, and should not be confused with a positive correlation. Rather, we can think of it in terms of "can predict", or "causality". Here, the findings show that application launches drive – or can predict – emotions, but the exact nature of this effect can in fact vary substantially over time as it is non-linear.

When interpreting the data points, we ignore the red dots as they are labelled by CCM as invalid: for these dots, CCM is not able to provide explanatory power that is greater than a plain correlation.

In our results, however, there are also some participants for whom emotion drives application launch (blue dots to the left of the graphs' origin). For example, contempt (1 participant), disgust (1 participant), joy (1 participant), sadness (3 participants), and surprise (2 participants) drive the number of launched applications, which means that once these participants experienced certain emotions, those emotions influenced the number of applications they launched.

Next, we repeat the analysis but this time considering the app launches across different application categories, to see if the causal relationship varies between categories. First, we present causality between Communication apps (*e.g.,* Messenger, WhatsApp, Telegram) and emotions in Figure 6 (Supplementary Materials), and observe that for most of the emotions (besides sadness) application launches drive emotions, with a stronger effect for surprise and contempt as compared to disgust and joy.

We also investigate the causal relationship between emotions and Social applications (*e.g.,* Facebook, Instagram) and present it in Figure 7 (Supplementary Materials), where we observe a strong effect of application launches on joy and contempt. However, the causal effect of application launch on disgust is not as strong when using social applications. Conversely, sadness and surprise drive application launches in this category.

We consider the causal relationship between Work applications (*e.g.,* Productivity, Education, and Business apps) and user emotions in Figure 8 (Supplementary Materials). We find that application launches drive most of the emotions, with a strong causal effect on contempt, disgust, and joy, as compared to sadness, and surprise. For this category of apps, user emotion does not drive app launches.

Finally, we examine causality between Entertainment applications (*e.g.,* games, YouTube, Netflix) and user emotions, which we visualise in Figure 9 (Supplementary Materials). We observe a strong causal effect of application launch on joy, as compared to contempt, disgust, and sadness. We can also observe that surprise drives application launches.

Application Use Duration vs. User Emotions

We also analyse the causality between emotions and app use *duration* (for all application categories), which we visualise in Figure 10 (Supplementary Materials), where we observe that for the majority of participants, application use duration drives emotions. However, there are some participants for whom certain emotions drive app use duration. For example, contempt drives app use duration for 1 participant, disgust drives app use for 4 participants, joy drives app use for 1 participant, and sadness and surprise drive application use duration for 3 and 2 participants respectively.

We also studied the causality between emotions and duration of different application *categories*. First, we visualise causality between Communication apps and emotions in Figure 11

Metric	Contempt	Disgust	Joy	Sadness	Surprise
Total Apps Usage Duration	0.091 (0.172)	0.047 (0.220)	0.138 (0.191)	0.086 (0.214)	0.144 (0.247)
Communication Apps Usage Duration	-0.118 (0.534)	0.131 (0.393)	0.124 (0.331)	0.122 (0.388)	0.177 (0.363)
Social Apps Usage Duration	0.221 (0.387)	0.093 (0.680)	0.129 (0.360)	0.168 (0.215)	-0.025 (0.469)
Work Apps Usage Duration	-0.060 (0.558)	0.064 (0.574)	-0.180 (0.316)	0.034 (0.547)	0.265 (0.315)
Entertainment Apps Usage Duration	0.149 (0.302)	-0.010 (0.616)	0.256 (0.471)	-0.050 (0.680)	-0.359 (0.379)

Table 2: Effect Sizes (and SD) for Causality between Application Usage Duration and Emotions. Blue: phone use drives emotions; Orange: emotion drives phone use.

(Supplementary Materials) – we can observe that for such emotions as disgust, joy, sadness, and surprise, duration drives emotions. Nevertheless, there exist several participants for whom emotions drive use duration of communication applications. Moreover, contempt drives use of communication apps for most participants.

Second, the results for social applications are shown in Figure 12 (Supplementary Materials) – we can observe that for most of the emotions (excluding surprise) use duration of social apps drives emotions. However, for certain participants, their emotions drive social app use. The plot does not show a strong causal effect between social apps use duration and surprise.

Third, the results for work applications are presented in Figure 13 (Supplementary Materials) – we can see that surprise is driven by application use duration; same as for disgust, and sadness, however the effect for disgust, and sadness is not very strong. In contrast, contempt and joy drive application use duration of work applications.

Finally, the results for entertainment apps are presented in Figure 14 (Supplementary Materials) – we can observe that entertainment applications use duration drive the experience of contempt and joy. Conversely, the experience of sadness and surprise drives the use of entertainment applications. We can also see that there is no causality relationship between disgust and use duration for entertainment apps, as the effect size value is close to zero for this emotion.

Qualitative Results

We used thematic analysis to analyse the data collected from our semi-structured interviews with participants. First, we read through each of the responses. Then three of the authors independently completed initial coding on the data. Next, the researchers compared initial codes and agreed on the final codes according to their similarity (*e.g., "I try to focus on something else"* and *"I get myself distracted"* would be merged to *"Diverting attention"*). We then independently coded participants' responses according to the finalised codes. Finally, we reviewed our coding and identified the themes described below.

Theme 1 – Content Matters

Whereas we categorised our data with regard to specific apps, many apps deliver variable content. The majority of our participants ($N = 23$) mentioned that their emotional states can be affected by the content of the applications, *e.g.,* positive content directs their emotions towards positive emotions, and negative content drives them towards negative emotions: *"It's the content of the app, for example videos or news, that can be sad or funny"* (P01), *"Depending on content of the app my emotions change"* (P08). Many participants ($N = 10$) mentioned that Social media and Communication apps in particular affect their emotions: *"Instagram or Facebook, depending on the content, make me feel either happy or sad"* (P06).

Theme 2 – App Use Drives Emotions

Some participants ($N = 5$) claim that particular apps drive their emotions towards certain directions. According to their comments, entertainment apps (*e.g.,* Reddit, Quora, YouTube, 9GAG, or music apps) cause positive emotions. For example: *"YouTube or humour sites (9GAG) make me feel happy"* (P04).

A couple of participants stated that social media apps lead to negative emotions as they lead participants to compare themselves to their friends on social media. *"Facebook and Instagram make you reflect on your life compared to others', Facebook is not really a happy app"* (P30).

Several participants ($N = 7$) mentioned that apps that connect them to their families lead to certain emotions depending on their personal situation and relation to the family or friends. Positive for some: *"WhatsApp and Instagram make me feel good as I talk to my close people"* (P28), and negative for others: *"Seeing my family on Facebook makes me sad as I feel excluded"* (P14).

Theme 3 – Emotions Drive App Use

We asked our participants if their emotions influence the choice of the application they use. Some participants ($N = 12$) said that when experiencing negative emotions, they seek content to feel uplifted: *"If sad, then I use web browser to find funny or interesting content to read to cheer me up"* (P26), *"If I am sad, then I use apps to make me happy (e.g., 9GAG)..."* (P16), whereas other participants seek for motivation: *"I use Pinterest to look up for ideas and motivation. For example, I am now into fitness so I look up motivations to do fitness, or I look up for motivation to do something creative"* (P28).

Some participants ($N = 5$) also reported being more active on social media when they are experiencing positive emotions: *"If I am happy I post more and use social media more"* (P15), partially to share their happiness: *"When happy, I use social media to share my happiness, or messaging apps"* (P13).

In contrast, many participants reported a decrease in phone use when they are experiencing positive emotions. Instead they prefer outdoor activities and spending time with friends ($N = 12$) or working ($N = 10$): *"When feeling positive emotions I don't do anything special, continue working or doing what I was doing, nothing on the smartphone"* (P20).

A few participants ($N = 6$) reported that when experiencing negative emotions, they seek distractions and some use their smartphones as a distraction. *"I try to focus on something else, or talk to other people. Phone can also be a distraction: call or text someone"* (P03). Using different apps on the phone is also seen as a distraction: *"I avoid things that make me feel negative, I try to distract myself by using Facebook, reading a book, talking to people. I play music on my phone (sad music)"* (P13). Moreover, one participant states that when feeling sad, they use Communication or Social Media applications to find support: *"If I am depressed or sad, I use Facebook or WeChat to seek support"* (P25).

Validating the Presence of Outliers

Overall, our results show that application use drives participants' emotions more than vice-versa. However, when looking at the detailed graphs produced by CCM, we find that there exist certain "outlier" participants: for them, emotions drive application use. We chose to conduct additional analysis for these participants, to investigate whether the quantitative results are reliable or can be attributed to "noise". We conducted our analysis by contrasting the CCM results to the comments these participants made during the interview. For instance, we observed that certain participants ($N = 3$) mentioned in the interviews that they prefer to do work when they experience positive emotions. When we look at their individual quantitative data, we found that they were the participants for whom joy drives use of work apps (*e.g.,* Productivity and Business apps) (Figure 13, Supplementary Materials).

We also investigated the causal relationship between sadness and Entertainment apps. We found that sadness drives entertainment app use for two participants (P07, P16). In their interview answers, they both stated that when experiencing negative emotions they listen to music (P07) and play games (P16) (Figure 14, Supplementary Materials). In addition, in Figure 7 (Supplementary Materials) we see that for two participants, sadness drives social application launch. We examine those participants' data individually and find that one of them (P25) mentioned in their interview that when feeling sad, they use social media applications to feel better. Same participant happens to be an "outlier" in a causal relationship between communication applications and their use duration (Figure 12, Supplementary Materials). These examples demonstrate that causal relationship between application use and user emotions can operate at an individual level. For some people, emotions drive application use; however for others, application use drives their emotions.

DISCUSSION

Understanding Smartphone Use and User Emotions

Our results suggest that the number of launched applications drives our participants' emotions for the majority of participants. This means that the more applications our participants launched, the more likely participants are to experience a wider range of emotions. This might be due to the fact that different applications lead to different emotions, based on the content of the application. This finding is also supported by our qualitative data, where participants mention that depending on the content of the application, they might experience positive or negative emotions. However, for some participants the causality is reversed: emotions drive the number of application launches. This finding can be explained by the fact that when experiencing certain negative emotions, people tend to find distractions and divert their attention [1, 93]. Moreover, this interpretation is supported by our qualitative data, as some of our participants mention this being the case for them. For example, when asked about their activities when experiencing negative emotions several participants mentioned they would seek distraction. Alternatively, a strong emotional experience, positive or negative, may motivate someone to share the experience using their phone. Therefore, the number of app launches might be an indicator of participants trying to divert their attention, or share their experience.

As for our investigation of the causal relationship between duration of application use and participant emotions, we found that for the majority of participants application use duration drives emotions. This finding is supported by our qualitative results, in which participants mentioned that application use drives emotions particularly when exposed for a longer duration: positive content can lead to positive emotions, whereas negative content tends to lead to negative emotions. We also found that for work applications, joy drives application use duration for the majority of participants. Our qualitative data supports the finding that when experiencing positive emotions, participants tend to continue working, which has also been showcased in prior literature [92].

However, we also observe in our results that for certain applications categories (*e.g.,* Entertainment, Social) the causal relationship between emotions and application use is stronger. For example, in the case of entertainment applications, the causal effect of applications use duration and joy or sadness is considerably stronger than the relationship between app use and other emotions. This is most likely due to the diverse content of entertainment applications that can lead to the user experiencing both joyful (*e.g.,* watching a comedy on Netflix) and distressing emotions (*e.g.,* losing in a game [42]). The same explanation can be applied to Social applications [5].

Emotion Detection and Analysis

In recent years there has been a rapid growth of research in the area of affective state sensing, particularly within the HCI community [107]. The potential benefits of successfully sensing affective states have been discussed in *e.g.,* Calvo *et al.* [14]. However, mood and emotional state sensing is challenging, especially when conducted in-the-wild, such as data collection issues. Therefore, investigating and developing effective mood and emotional state detection techniques using accessible ubiquitous and wearable technology, without relying on high-cost bespoke hardware [107], is an important challenge. Although self-report methods such as the Experience Sampling Method have seen increased popularity in our community, automated sensing techniques have a number of potential advantages. First, we can collect many more samples throughout the day – providing a more detailed report on the user's affective state. Second, we do not need to interrupt our study participants. Third, previous work shows that the reliability of self-report data deteriorates after several weeks [96], whereas automated methods should maintain a consistent quality level over long periods of time. Today's smartphones are widely used in people's day-to-day lives for various purposes and are equipped with many sensors (*e.g.,* accelerometer, gyroscope). As such, they can act as a scientific tool to collect contextual information on user's emotional, social and physical behaviour.

In such work it is important to not only accurately detect emotions, but also to reliably analyse the relationships between emotions and other phenomena. In this work we use the Convergent Cross Mapping method, as it has previously been shown to best identify causality (rather than merely correlation) between two variables [89]. The method is applicable to our work, as we investigate the causal relationship between users' emotional state and application use (number of launches and duration), and have very granular data. Our results show that for some participants application use causes them to experience certain emotions, while for other participants their emotional states influence their smartphone use. These findings are in line with our qualitative data, as our participants mention that certain application use leads them to feel certain ways, and that the way they feel plays a role on deciding which applications to use. Furthermore, recent work has utilised CCM to determine the leader and the follower of the facial expressions between the two interlocutors. The results of the study reveal the presence of bidirectional causal couplings of the facial dynamics; hence, CCM can establish evidence for causal behavioural interactions [75].

Finally, it is important to consider the underlying reliability of our emotional data. As facial expressions are commonly used to detect emotional states [18], we used the Affectiva SDK to collect the user's emotional state from their facial expressions. We acknowledge that there is an ongoing debate in the research community that facial expressions may not reflect people's true emotional feelings [29]. Nevertheless, we validated the reliability of emotions derived from facial expression in numerous ways to ensure the robustness of our findings. First, there is a statistically significant positive correlation between the self-reports and the Affectiva valence data, even though both are measured at different frequencies. Second, we plotted the weekly trends of joy and sadness, which are in line with previous work [88], in that towards the weekend participants' joy increases while it drops during the week days; whereas sadness drops towards the weekend, and is at its peak on week days. Third, our outlier analysis confirms the agreement between our quantitative and qualitative data from participants that reported using certain apps when they were feeling a certain way. Finally, the results of our validation study show a high level of reliability in detecting surprise, joy, disgust, sadness and contempt. However, the accuracy of Affectiva in detecting anger and fear was considerably lower. Therefore, we removed these emotions from our analysis.

Implications for Research and Practice

One of the objectives of HCI, and in particular UbiComp research, is to surround humans with sensitive and responsive technology, unobtrusively embedded in the environment [102]. To sense and respond to human needs, the environment should be able to invigilate humans' physical, mental, and emotional state [11]. Previous work has shown that smartphones can be successfully leveraged to detect physical context (*e.g.,* environment [2, 82], physical [103] and situational impairments [34, 83]) as well as to the cognitive context of the user, such as their emotional state [8, 61] or stress [81].

Although prior work has studied the effect of smartphone use on a user's emotional state [61, 65], in our work we show that the causal relationship between emotions and application use is in fact bidirectional. Furthermore, our findings demonstrate that the causal relationship between emotional state and smartphone use varies between people. This is unsurprising, as emotions and behaviour are intrinsically personal. What remains unclear from our analysis is whether the stark differences between participants are ephemeral or persistent. In other words, do participants experience causality in the same direction over long period of times, or does causality reverse over time. For instance, it may be plausible to expect that under stress, the direction of the causality may change.

It is also important to consider the implications of our findings that emotional states can drive application use, as a great deal of HCI work promotes emotional well-being [3, 9, 15]. If someone's experience of a particular emotional state leads to a certain type of smartphone use which can in turn influence emotional states, it may be that the person is using their phone in an attempt to influence their emotional state. "Emotion regulation" [38] has been defined as: *"all of the conscious and non-conscious strategies we use to increase, maintain, or decrease one or more components of an emotional response"*. A large body of work in psychology has demonstrated the importance of emotion regulation in daily life and its important influence on well-being [39].

Some emotion regulation strategies that have been studied by psychologists, such as distraction, may be particularly suited to technological mediation. Our findings suggest that the investigation of *digital* emotion regulation using smartphones should be conducted at the level of individual users, since different participants exhibited different relationships between emotion and phone use. Personalised predictive models would produce more accurate results in detecting emotion regulation for particular users, as those are built and trained on a single user's data. Such an approach could also extend existing smartphone features designed to help people understand their phone use. For example, Apple's iOS now allows users to track how much time they spend on applications. Personalised models that incorporate a user's emotional state during phone use could help users understand how their phone use may be driven by their emotions and vice-versa. This could allow users to better understand the emotional causes and consequences of their phone use habits and, if appropriate, try to change them.

Limitations

Our study has a number of limitations. We lost 5.5% of the emotions data due to Affectiva not being able to detect emotions when the participants' face was not in full view of their smartphone's frontal camera. This is a generalised limitation and not a participant specific issue. This is an identified problem and was recently brought up by Khamis *et al.* [46]. Although the authors suggest instructing participants to hold the phone in specific ways, we did not follow this recommendation, in order to allow more realistic phone use conditions. Furthermore, we ran the study for an extensive period of time to ensure sufficient data points were collected. Moreover, we recognise that Affectiva performs better on high-quality images when compared to lower quality images that are more likely to occur in a real-world scenario. Therefore, we went beyond a laboratory validation by inclusion of ESMs and analysis of daily/weekly trends.

In addition, there is an ongoing debate about the CCM method within the research community. For example, Paluš and colleagues argue that CCM cannot be used to infer the direction of causality, but rather capture forecasting ability. In particular, they state that CCM lacks any arrow of time in its formulation [69]. Nevertheless, Sugihara *et al.* state that if variables are mutually coupled, they will cross map in both directions. Moreover, the authors state that the strength of coupling defines the relative strength of causal effect [89].

Furthermore, due to emerging challenges related to battery conservation and privacy protection, Android Accessibility services were switched off from time to time, which resulted in application data not being collected. In some cases our participants switched on the Accessibility services instantly when they saw the notification from the 'ApplicationTracker', but in some cases they did not. This resulted in data loss on smartphone use. Finally, some participants noticed a slight increase in battery use due to constant data collection during phone use. Although no participants reported a major complaint about battery drain, this issue could be mitigated by lowering the scanning rate.

CONCLUSION

In this work, we investigate the bidirectional causal relationship between the emotional state of users and their phone use. Our results show that for some participants, the use of particular apps causes them to experience certain emotions; however, for other participants, their emotions drive app use behaviour. We also found noticeable differences between different application categories with regards to their causal relationship with the users' emotional state. These quantitative findings corresponded with qualitative results showing that participants mentioned that certain applications cause them to experience certain emotions, and that feeling certain emotions led them to use certain applications. Our findings are a step towards building personalised models which can help users better understand the relationship between their phone use is intertwined with their emotional states. This could potentially lead to more effective decision-making with regards to smartphone use as well as better technology-mediated support for emotion regulation.

ACKNOWLEDGEMENTS

This work is partially funded by ARC Discovery Project DP190102627 and NHMRC grant 1170937.

REFERENCES

[1] Amelia Aldao, Susan Nolen-Hoeksema, and Susanne Schweizer. 2010. Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical psychology review* 30, 2 (2010), 217–237.

[2] Siamak Aram, Amedeo Troiano, and Eros Pasero. 2012. Environment sensing using smartphone. In *2012 IEEE Sensors Applications Symposium Proceedings*. IEEE, 1–4.

[3] Amid Ayobi, Paul Marshall, Anna L. Cox, and Yunan Chen. 2017. Quantifying the body and caring for the mind: self-tracking in multiple sclerosis. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 6889–6901.

[4] Jeremy N. Bailenson, Nick Yee, Scott Brave, Dan Merget, and David Koslow. 2007. Virtual interpersonal touch: expressing and recognizing emotions through haptic devices. *Human–Computer Interaction* 22, 3 (2007), 325–353.

[5] Eytan Bakshy, Solomon Messing, and Lada A. Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.

[6] C Daniel Batson, Laura L Shaw, and Kathryn C Oleson. 1992. Differentiating affect, mood, and emotion: toward functionally based conceptual distinctions. (1992).

[7] Christopher Beedie, Peter Terry, and Andrew Lane. 2005. Distinctions between emotion and mood. *Cognition and Emotion* 19, 6 (2005), 847–878. DOI: http://dx.doi.org/10.1080/02699930541000057

[8] Adriana Bianchi and James G. Phillips. 2005. Psychological predictors of problem mobile phone use. *CyberPsychology & Behavior* 8, 1 (2005), 39–51.

[9] Nataly Birbeck, Shaun Lawson, Kellie Morrissey, Tim Rapley, and Patrick Olivier. 2017. Self Harmony: rethinking hackathons to design and critique digital technologies for those affected by self-harm. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 146–157.

[10] Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. 2011. Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage. In *Proceedings of the 13th international conference on Human computer interaction with mobile devices and services*. ACM, 47–56.

[11] Tibor Bosse and Frank P.J. De Lange. 2008. Estimating emotion regulation capabilities. In *Proceedings of the 1st international conference on PErvasive Technologies Related to Assistive Environments*. ACM, 93.

[12] Michelle Nicole Burns, Mark Begale, Jennifer Duffecy, Darren Gergle, Chris J. Karr, Emily Giangrande, and David C. Mohr. 2011. Harnessing context sensing to develop a mobile intervention for depression. *Journal of medical Internet research* 13, 3 (2011).

[13] Paul Cairns, Pratyush Pandab, and Christopher Power. 2014. The influence of emotion on number entry errors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2293–2296.

[14] Rafael A. Calvo, Sidney D'Mello, Jonathan Matthew Gratch, and Arvid Kappas. 2015. *The Oxford handbook of affective computing*. Oxford University Press, USA.

[15] Scott A. Cambo, Daniel Avrahami, and Matthew L. Lee. 2017. BreakSense: Combining physiological and location sensing to promote mobility during work-breaks. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3595–3607.

[16] John M. Carroll. 1997. Human-computer interaction: psychology as a science of design. *Annual review of psychology* 48, 1 (1997), 61–83.

[17] Keng-hao Chang, Drew Fisher, John Canny, and Björn Hartmann. 2011. How's my mood and stress?: an efficient speech analysis library for unobtrusive monitoring on mobile phones. In *Proceedings of the 6th International Conference on Body Area Networks*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 71–77.

[18] Weixuan Chen and Rosalind W. Picard. 2017. Eliminating physiological information from facial videos. In *Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on*. IEEE, 48–55.

[19] Erika Chuang and Christoph Bregler. 2005. Mood swings: expressive speech animation. *ACM Transactions on Graphics (TOG)* 24, 2 (2005), 331–347.

[20] Emily Collins and Anna L. Cox. 2014. Switch on to games: Can digital games aid post-work recovery? *International Journal of Human-Computer Studies* 72, 8-9 (2014), 654–662.

[21] Stephen A. Coombes, Christopher M. Janelle, and Aaron R. Duley. 2005. Emotion and motor control: Movement attributes following affective picture processing. *Journal of motor behavior* 37, 6 (2005), 425–436.

[22] Mihaly Csikszentmihalyi and Reed Larson. 2014. *Validity and Reliability of the Experience-Sampling Method*. Springer Netherlands, Dordrecht, 35–54. DOI: http://dx.doi.org/10.1007/978-94-017-9088-8_3

[23] Raymond J. Dolan. 2002. Emotion, cognition, and behavior. *Science* 298, 5596 (2002), 1191–1194.

[24] Paul Ekman and Wallace V. Friesen. 1976. Measuring facial movement. *Environmental psychology and nonverbal behavior* 1, 1 (1976), 56–75.

[25] Paul Ekman and Dacher Keltner. 1997. Universal facial expressions of emotion. *Segerstrale U, P. Molnar P, eds. Nonverbal communication: Where nature meets culture* (1997), 27–46.

[26] Clayton Epp, Michael Lippold, and Regan L. Mandryk. 2011. Identifying emotional states using keystroke dynamics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 715–724.

[27] Hossein Falaki, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. 2010. Diversity in smartphone usage. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 179–194.

[28] Denzil Ferreira, Vassilis Kostakos, and Anind K. Dey. 2015. AWARE: mobile context instrumentation framework. *Frontiers in ICT* 2 (2015), 6.

[29] Alan J. Fridlund. 2014. *Human facial expression: An evolutionary view*. Academic Press.

[30] E. Friesen and Paul Ekman. 1978. Facial action coding system: a technique for the measurement of facial movement. *Palo Alto* 3 (1978).

[31] Nico H. Frijda. 1986. *The emotions*. Cambridge University Press.

[32] Yuan Gao, Nadia Bianchi-Berthouze, and Hongying Meng. 2012. What does touch tell us about emotions in touchscreen-based gameplay? *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 4 (2012), 31.

[33] William W. Gaver. 1991. Technology Affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)*. ACM, New York, NY, USA, 79–84. DOI: http://dx.doi.org/10.1145/108844.108856

[34] Mayank Goel, Leah Findlater, and Jacob O. Wobbrock. 2012. WalkType: Using Accelerometer Data to Accomodate Situational Impairments in Mobile Touch Screen Text Entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 2687–2696. DOI: http://dx.doi.org/10.1145/2207676.2208662

[35] Jorge Goncalves, Pratyush Pandab, Denzil Ferreira, Mohammad Ghahramani, Guoying Zhao, and Vassilis Kostakos. 2014. Projective testing of diurnal collective emotion. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 487–497.

[36] Kendall Goodrich and Marieke De Mooij. 2014. How 'social' are social media? A cross-cultural comparison of online and offline purchase decision influences. *Journal of Marketing Communications* 20, 1-2 (2014), 103–116.

[37] C.W.J. Granger. 1988. Causality, cointegration, and control. *Journal of Economic Dynamics and Control* 12, 2 (1988), 551 – 559. DOI: http://dx.doi.org/https://doi.org/10.1016/0165-1889(88)90055-3

[38] James J. Gross. 1998. The emerging field of emotion regulation: An integrative review. *Review of general psychology* 2, 3 (1998), 271–299.

[39] James J. Gross. 2015. Emotion regulation: Current status and future prospects. *Psychological Inquiry* 26, 1 (2015), 1–26.

[40] James J. Gross and Oliver P. John. 2003. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology* 85, 2 (2003), 348.

[41] Bjarte Hannisdal and Lee Hsiang Liow. 2018. Causality from palaeontological time series. *Palaeontology* 61, 4 (2018), 495–509. DOI: http://dx.doi.org/10.1111/pala.12370

[42] Richard L. Hazlett. 2006. Measuring Emotional Valence During Interactive Experiences: Boys at Video Game Play. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 1023–1026. DOI: http://dx.doi.org/10.1145/1124772.1124925

[43] Chin-Lung Hsu and Hsi-Peng Lu. 2004. Why do people play on-line games? An extended TAM with social influences and flow experience. *Information Management* 41, 7 (2004), 853 – 868. DOI: http://dx.doi.org/https://doi.org/10.1016/j.im.2003.08.014

[44] Galen Chin-Lun Hung, Pei-Ching Yang, Chia-Chi Chang, Jung-Hsien Chiang, and Ying-Yeh Chen. 2016. Predicting negative emotions based on mobile phone usage patterns: an exploratory study. *JMIR research protocols* 5, 3 (2016).

[45] Alice M. Isen, Andrew S. Rosenzweig, and Mark J. Young. 1991. The influence of positive affect on clinical problem solving. *Medical Decision Making* 11, 3 (1991), 221–227.

[46] Mohamed Khamis, Anita Baier, Niels Henze, Florian Alt, and Andreas Bulling. 2018. Understanding Face and Eye Visibility in Front-Facing Cameras of Smartphones used in the Wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 280.

[47] Preeti Khanna and Mukundan Sasikumar. 2010. Recognising emotions from keyboard stroke pattern. *International journal of computer applications* 11, 9 (2010), 1–5.

[48] Adam D. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24 (2014), 8788–8790.

[49] Amanda E. Krause, Adrian C. North, and Lauren Y. Hewitt. 2015. Music-listening in everyday life: Devices and choice. *Psychology of music* 43, 2 (2015), 155–170.

[50] Peter J. Lang, Mark K. Greenwald, Margaret M. Bradley, and Alfons O. Hamm. 1993. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 3 (1993), 261–273.

[51] Oliver Langner, Ron Dotsch, Gijsbert Bijlstra, Daniel H.J. Wigboldus, Skyler T. Hawk, and Ad Van Knippenberg. 2010. Presentation and validation of the Radboud Faces Database. *Cognition and emotion* 24, 8 (2010), 1377–1388.

[52] Reed Larson, Bernard Hecker, and Julie Norem. 1985. Students' experience with research projects: Pains, enjoyment and success. *The High School Journal* 69, 1 (1985), 61–69.

[53] Richard S. Lazarus. 1991. *Emotion and adaptation*. Oxford University Press on Demand.

[54] Hosub Lee, Young Sang Choi, Sunjae Lee, and I.P. Park. 2012. Towards unobtrusive emotion recognition for affective social communication. In *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*. IEEE, 260–264.

[55] Peter Leijdekkers, Valerie Gay, and Frederick Wong. 2013. CaptureMyEmotion: A mobile app to improve emotion learning for autistic children using sensors. In *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*. IEEE, 381–384.

[56] Robert LiKamWa, Yunxin Liu, Nicholas D. Lane, and Lin Zhong. 2013. Moodscope: Building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. ACM, 389–402.

[57] Manas K. Mandal, Rakesh Pandey, and Akhouri B. Prasad. 1998. Facial expressions of emotions and schizophrenia: a review. *Schizophrenia bulletin* 24, 3 (1998), 399–412.

[58] John McCarthy and Peter Wright. 2004. Technology As Experience. *Interactions* 11, 5 (Sept. 2004), 42–43. DOI: http://dx.doi.org/10.1145/1015530.1015549

[59] Daniel McDuff, Abdelrahman Mahmoud, Mohammad Mavadati, May Amr, Jay Turcot, and Rana el Kaliouby. 2016. AFFDEX SDK: A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 3723–3726. DOI:http://dx.doi.org/10.1145/2851581.2890247

[60] Abhinav Mehrotra, Sandrine R. Müller, Gabriella M. Harari, Samuel D. Gosling, Cecilia Mascolo, Mirco Musolesi, and Peter J. Rentfrow. 2017a. Understanding the role of places and activities on mobile phone interaction and usage patterns. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 84.

[61] Abhinav Mehrotra, Fani Tsapeli, Robert Hendley, and Mirco Musolesi. 2017b. MyTraces: investigating correlation and causation between users' emotional states and mobile phone interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 83.

[62] Alistair Morrison, Xiaoyu Xiong, Matthew Higgs, Marek Bell, and Matthew Chalmers. 2018. A Large-Scale Study of iPhone App Launch Behaviour. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 344.

[63] Aske Mottelson and Kasper Hornbæk. 2016. An affect detection technique using mobile commodity sensors in the wild. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 781–792.

[64] Jessica Gall Myrick. 2015. Emotion regulation, procrastination, and watching cat videos online: Who watches Internet cats, why, and to what effect? *Computers in human behavior* 52 (2015), 168–176.

[65] Evangelos Niforatos and Evangelos Karapanos. 2015. EmoSnaps: a mobile application for emotion recall from facial expressions. *Personal and Ubiquitous Computing* 19, 2 (2015), 425–444.

[66] Donald A. Norman. 1986. *Cognitive Engineering*. Lawrence Erlbaum Associates, 31–61.

[67] Donald A. Norman. 2004. *Emotional design: Why we love (or hate) everyday things*. Basic Civitas Books.

[68] Michal Olszanowski, Grzegorz Pochwatko, Krzysztof Kuklinski, Michal Scibor-Rylski, Peter Lewinski, and Rafal K. Ohme. 2015. Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. *Frontiers in psychology* 5 (2015), 1516.

[69] Milan Paluš, Anna Krakovská, Jozef Jakubík, and Martina Chvosteková. 2018. Causality, dynamical systems and the arrow of time. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28, 7 (2018), 075307.

[70] Reinhard Pekrun, Thomas Goetz, Wolfram Titz, and Raymond P. Perry. 2002. Positive emotions in education. (2002).

[71] Ella Peltonen, Eemil Lagerspetz, Jonatan Hamberg, Abhinav Mehrotra, Mirco Musolesi, Petteri Nurmi, and Sasu Tarkoma. 2018. The hidden image of mobile apps: geographic, demographic, and cultural factors in mobile usage. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 10.

[72] Rosalind W. Picard. 2009. Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3575–3584.

[73] Ming-Zher Poh, Daniel J. McDuff, and Rosalind W. Picard. 2011. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE transactions on biomedical engineering* 58, 1 (2011), 7–11.

[74] John P. Pollak, Phil Adams, and Geri Gay. 2011. PAM: a photographic affect meter for frequent, in situ measurement of affect. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 725–734.

[75] Eric Postma and Marie Nilsenova. 2016. Measuring the causal dynamics of facial interaction. In *Proceedings of the Cognitive Science Conference*.

[76] Michael J. Power and Tim Dalgleish. 1999. *Handbook of cognition and emotion*. Wiley.

[77] Mashfiqui Rabbi, Shahid Ali, Tanzeem Choudhury, and Ethan Berke. 2011. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 385–394.

[78] Johnmarshall Reeve. 2014. *Understanding motivation and emotion*. John Wiley & Sons.

[79] James A. Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.

[80] Zhanna Sarsenbayeva, Denzil Ferreira, Niels van Berkel, Chu Luo, Mikko Vaisanen, Vassilis Kostakos, and Jorge Goncalves. 2017. Vision-based Happiness Inference: A Feasibility Case-study. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17)*. ACM, New York, NY, USA, 494–499. DOI:http://dx.doi.org/10.1145/3123024.3124438

[81] Zhanna Sarsenbayeva, Niels van Berkel, Danula Hettiachchi, Weiwei Jiang, Tilman Dingler, Eduardo Velloso, Vassilis Kostakos, and Jorge Goncalves. 2019. Measuring the Effects of Stress on Mobile Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 24 (March 2019), 18 pages. DOI: http://dx.doi.org/10.1145/3314411

[82] Zhanna Sarsenbayeva, Niels van Berkel, Eduardo Velloso, Vassilis Kostakos, and Jorge Goncalves. 2018. Effect of Distinct Ambient Noise Types on Mobile Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 2, Article 82 (July 2018), 23 pages. DOI:http://dx.doi.org/10.1145/3214285

[83] Zhanna Sarsenbayeva, Niels van Berkel, Aku Visuri, Sirkka Rissanen, Hannu Rintamäki, Vassilis Kostakos, and Jorge Goncalves. 2017. Sensing Cold-Induced Situational Impairments in Mobile Interaction Using Battery Temperature. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 98 (Sept. 2017), 9 pages. DOI:http://dx.doi.org/10.1145/3130963

[84] Nikolaos Savva and Nadia Bianchi-Berthouze. 2011. Automatic recognition of affective body movement in a video game scenario. In *International Conference on Intelligent Technologies for interactive entertainment*. Springer, 149–159.

[85] Thiago H. Silva, Pedro O.S. Vaz de Melo, Jussara M. Almeida, Mirco Musolesi, and Antonio A.F. Loureiro. 2014. You are what you eat (and drink): Identifying cultural boundaries by analyzing food and drink habits in foursquare. In *Eighth International AAAI Conference on Weblogs and Social Media*.

[86] Wook Song, Yeseong Kim, Hakbong Kim, Jehun Lim, and Jihong Kim. 2014. Personalized optimization for android smartphones. *ACM Transactions on Embedded Computing Systems (TECS)* 13, 2s (2014), 60.

[87] Sabrina Stöckli, Michael Schulte-Mecklenbeck, Stefan Borer, and Andrea C. Samson. 2018. Facial expression analysis with AFFDEX and FACET: A validation study. *Behavior research methods* 50, 4 (2018), 1446–1460.

[88] Arthur A. Stone, Stefan Schneider, and James K. Harter. 2012. Day-of-week mood patterns in the United States: On the existence of 'Blue Monday','Thank God it's Friday' and weekend effects. *The Journal of Positive Psychology* 7, 4 (2012), 306–314.

[89] George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. 2012. Detecting causality in complex ecosystems. *Science* 338, 6106 (2012), 496–500.

[90] Boyuan Sun, Qiang Ma, Shanfeng Zhang, Kebin Liu, and Yunhao Liu. 2017. iSelf: Towards Cold-Start Emotion Labeling Using Transfer Learning with Smartphones. *ACM Trans. Sen. Netw.* 13, 4, Article 30 (Sept. 2017), 22 pages. DOI: http://dx.doi.org/10.1145/3121049

[91] Floris Takens. 1981. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980*, David Rand and Lai-Sang Young (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 366–381.

[92] Maxime Taquet, Jordi Quoidbach, Yves-Alexandre De Montjoye, Martin Desseilles, and James J. Gross. 2016. Hedonism and the choice of everyday activities. *Proceedings of the National Academy of Sciences* 113, 35 (2016), 9769–9773.

[93] Robert E. Thayer, J. Robert Newman, and Tracey M. McClain. 1994. Self-regulation of mood: Strategies for changing a bad mood, raising energy, and reducing tension. *Journal of personality and social psychology* 67, 5 (1994), 910.

[94] Lucia R. Valmaggia, Leila Latif, Matthew J. Kempton, and Maria Rus-Calafell. 2016. Virtual reality in the psychological treatment for mental health problems: an systematic review of recent evidence. *Psychiatry Research* 236 (2016), 189–195.

[95] Niels van Berkel, Simon Dennis, Michael Zyphur, Jinjing Li, Andrew Heathcote, and Vassilis Kostakos. 2020. Modelling Interaction as a Complex System. *Human-Computer Interaction* (2020), to appear.

[96] Niels van Berkel, Jorge Goncalves, Peter Koval, Simo Hosio, Tilman Dingler, Denzil Ferreira, and Vassilis Kostakos. 2019. Context-Informed Scheduling and Analysis: Improving Accuracy of Mobile Self-Reports. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3290605.3300281

[97] Job Van Der Schalk, Skyler T. Hawk, Agneta H. Fischer, and Bertjan Doosje. 2011. Moving faces, looking places: validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion* 11, 4 (2011), 907.

[98] Egbert H Van Nes, Marten Scheffer, Victor Brovkin, Timothy M Lenton, Hao Ye, Ethan Deyle, and George Sugihara. 2015. Causal feedbacks in climate change. *Nature Climate Change* 5, 5 (2015), 445.

[99] Aku Visuri, Zhanna Sarsenbayeva, Jorge Goncalves, Evangelos Karapanos, and Simon Jones. 2016. Impact of mood changes on application selection. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 535–540.

[100] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T. Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 3–14.

[101] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54, 6 (1988), 1063.

[102] Mark Weiser. 1993. Ubiquitous computing. *Computer* 10 (1993), 71–72.

[103] Jacob O. Wobbrock, Shaun K. Kane, Krzysztof Z. Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)* 3, 3 (2011), 9.

[104] Hao Ye, Ethan R Deyle, Luis J Gilarranz, and George Sugihara. 2015. Distinguishing time-delayed causal interactions using convergent cross mapping. *Scientific reports* 5 (2015), 14750.

[105] Hao Ye and et al. 2016. Package rEDM. (2016). https://cran.r-project.org/web/packages/rEDM/

[106] Ligang Zhang and Dian Tjondronegoro. 2011. Facial expression recognition using facial movement features. *IEEE Transactions on Affective Computing* 2, 4 (2011), 219–229.

[107] Xiao Zhang, Wenzhong Li, Xu Chen, and Sanglu Lu. 2018. MoodExplorer: Towards Compound Emotion Detection via Smartphone Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 176 (Jan. 2018), 30 pages. DOI: http://dx.doi.org/10.1145/3161414