"Living Within Four Walls": Exploring Emotional and Social Dynamics in Mobile Usage During Home Confinement

NAN GAO*, Department of Computer Science and Technology Tsinghua University, China and University of New South Wales (UNSW), Australia SAM NOLAN*, RMIT University, Australia KAIXIN JI*, RMIT University, Australia SHAKILA KHAN RUMI, RMIT University, Australia JUDITH SIMONE HEINISCH, University of Kassel, Germany CHRISTOPH ANDERSON, University of Kassel, Germany KLAUS DAVID, University of Kassel, Germany FLORA D. SALIM, University of New South Wales (UNSW), Australia

Home confinement, a situation experienced by individuals for reasons ranging from medical quarantines, rehabilitation needs, disability accommodations, and remote working, is a common yet impactful aspect of modern life. While essential in various scenarios, confinement within the home environment can profoundly influence mental well-being and digital device usage. Using the COVID-19 lockdown as a case study, this research explores the emotional and social effects of prolonged home confinement on mobile device usage. We conducted an in-situ study with 32 participants, analyzing three weeks of mobile usage data to assess emotional well-being and social dynamics in restricted environments. Our findings reveal that app usage patterns serve as strong indicators of emotional states, offering insights into how digital interactions can reflect and influence well-being during isolation. This study highlights the potential for developing targeted interventions and support systems for individuals in long-term home confinement, including those with chronic illness, recovery needs, or permanent remote work situations.

CCS Concepts: • Human-centered computing → Ubiquitous and mobile computing; • Applied computing;

Additional Key Words and Phrases: Emotion, Home Confinement, Mobile Sensing, COVID-19 Lockdown, Social Roles, Well-Being

ACM Reference Format:

Nan Gao, Sam Nolan, Kaixin Ji, Shakila Khan Rumi, Judith Simone Heinisch, Christoph Anderson, Klaus David, and Flora D. Salim. 2024. "Living Within Four Walls": Exploring Emotional and Social Dynamics in Mobile Usage During Home Confinement. 1, 1 (December 2024), 26 pages. https://doi.org/10.1145/1122445.XXXX

*Equally contributing first authors.

Authors' addresses: Nan Gao, Department of Computer Science and Technology, Tsinghua University, Beijing, China, 100081, University of New South Wales (UNSW), Sydney, Australia, 1466, nangao@tsinghua.edu.cn; Sam Nolan, samnolan555@gmail.com, RMIT University, Melbourne, Australia, 3000; Kaixin Ji, kaixin.ji@student.rmit.edu.au, RMIT University, Melbourne, Australia, 3000; Shakila Khan Rumi, shakilakhanrumi@gmail.com, RMIT University, Melbourne, Australia, 3000; Judith Simone Heinisch, judith.heinisch@comtec.eecs.uni-kassel.de, University of Kassel, Kassel, Germany, 34131; Christoph Anderson, c_anderson@posteo.de, University of Kassel, Kassel, Germany, 34131; Klaus David, david@uni-kassel.de, University of Kassel, Kassel, Germany, 34131; Flora D. Salim, flora.salim@unsw.edu.au, University of New South Wales (UNSW), Sydney, Australia, 1466.

© 2024 Association for Computing Machinery.

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

1 INTRODUCTION

Home confinement, defined as the restriction of movements within the home [51] due to illness, recovery, disability, or other personal circumstances, is a common phenomenon that everyone may experience during their lives [10, 37]. Research has shown that such confinement can adversely affect mental well-being, leading to stress, anxiety, and depression [5, 65]. For instance, Bartoszek et al. [5] found that quarantined individuals experienced heightened levels feelings of depression, insomnia, loneliness, and fatigue. Understanding these impacts is essential for developing effective targeted interventions to support individuals during such periods.

Traditionally, mental well-being during home confinement has been assessed using survey-based instruments [21, 65], which can be time-consuming, labour-intensive, and prone to biases such as social desirability and reflection bias [22, 58]. Recent advances in mobile sensing, however, offer a more objective, unobtrusive and real-time alternative [14, 60]. While methods relying on physical movement data (e.g., GPS and accelerometers) have been widely used for mental wellbeing modelling such as social anxiety, affect, depression and social isolation [14, 60], they are less effective in the context of indoor confinement, where movement patterns (e.g., step counts, visited locations) are limited [60]. This study shifts focus to more granular data, such as app usage patterns, to provide deeper insights into emotional wellbeing during confinement.

The diversity of individual circumstances complicates the study of home confinement. Factors such as the reason for confinement, personal health conditions, and social environment can vary greatly, complicating the creation of a standardized model. However, the COVID-19 lockdown provided a unique, global and natural context in which the majority of individuals experienced similar restrictions, with a marked increased reliance on digital technologies [56, 69]. This shared experience allows for a more uniform analysis of how mobile device usage reflects emotional states, social roles, and work-life dynamics during such periods. Espeically, while this study uses the lockdown as a descriptive case study [86], the insights gained may also be applicable to other forms of home confinement, such as those experienced by individuals with chronic illnesses, in post-recovery rehabilitation, or those adapting to permanent remote work arrangements.

In this study, we explore the relationship between mobile device interactions, emotional well-being, and social roles, during COVID-19 lockdown, as a case study for home confinement. We focus on two key aspects of well-being: emotions and social roles, as both are critical to mental health [16, 53]. By utilizing the *Valence* and *Arousal* dimensions from the *Circumplex Model of Affect* [66] and *Social Role* from the *Role Theory* [7], we aim to provide a comprehensive understanding of how mobile device usage during lockdown reflects emotional and social dynamics. Our research seeks to address the following questions: **RQ1**. *How are emotions and social roles during COVID-19 home confinement perceived through self-reports and inferred from app usage patterns?* **RQ2**. *Can mobile usage data be used to model emotional well-being during home confinement, and which features are most predictive?*

To answer these questions, we conducted a three-week data collection during the COVID-19 lockdown in Melbourne, Australia, involving 32 participants. We analyzed perceived emotional states, social roles, and work-life balance through *Experience Sampling Method* (ESM) [6] and end-of-day surveys. We also proposed a novel method, *App-based Information Gain Temporal Segmentation* (AIGTS), to extract user activities from app usage records by identifying segments and transition times between apps during the day. The resulting insights offer a deeper understanding of how app usage during confinement reflects emotional well-being and social dynamics, with implications for developing

Living Within Four Walls

digital interventions to support individuals in long-term home confinement. All scripts for this study are open source.¹ In summary, our contributions are as follows:

- We collected a rich dataset from 32 participants during a 3-week COVID-19 home confinement period in Melbourne, Australia. This dataset includes 502,485 records of users smartphone and desktop usage data, 1,749 ESM responses, and 265 end-of-day surveys. It provides a unique, real-world snapshot of emotional well-being and social dynamics during lockdown.
- We analyzed perceived valence, arousal, and social roles, revealing that the lockdown significantly impacted well-being, reducing emotional valence and increasing work hours due to interruptions or blurred work-life boundaries. These findings shed light on the complex emotional landscape during prolonged confinement.
- We introduced the App-based Information Gain Temporal Segmentation (AIGTS) method, an unsupervised technique for extracting user activities (e.g., video-watching, socializing, deep work) from app usage data. This method autonomously identified 16 distinct user activities, providing granular insights into how digital behaviours relate to emotional well-being.
- We demonstrated that mobile usage patterns can predict emotional well-being during home confinement. Our
 extensive experiments revealed key features that are most predictive of emotional states, offering new avenues
 for developing digital interventions to support individuals during isolation.

2 RELATED WORK

2.1 Impact of the Home Confinement on Mental Well-Being

Home confinement, whether due to edical conditions, disability, or societal events like pandemics, can significantly affect mental well-being. Prolonged isolation, particularly when coupled with limited physical activity and disrupted social interactions, is linked to increased risks of anxiety, depression, stress, and loneliness [9, 13, 75]. The psychological impact is often compounded by the loss of routine and control, leading to emotional disturbances such as irritability, confusion, and anger [38, 63]. These effects may persist even after restrictions are lifted [30].

The COVID-19 pandemic underscored these effects on a global scale. The World Health Organization [83] reported that the psychological toll of lockdowns and social distancing measures often surpassed the physical health risks. Studies during the pandemic have documented a significant increases in mental health issues, with depression and anxiety becoming widespread [15, 36, 84]. For example, Khubchandani et al. [35] found depression (39%) and anxiety (42%) surged in the U.S. post-lockdown. Similarly, Fountoulakis et al. [19] reported increased anxiety (45%) and suicidal thoughts (10.4%) in Greece. Loneliness was also a major issue, especially among those in long-term isolation [36].

Despite extensive research on the mental well-being impacts of various forms of home confinement, much of it is geographically limited or reflects different phases of isolation. This variability makes it challenging to fully understand the broader, long-term psychological effects of home confinement in diverse contexts. To address this gap, **RQ1** examines how emotions and social roles are experienced during COVID-19 home confinement, using self-reports and app usage patterns to infer these dynamics.

2.2 Digital Device Usage During Home Confinement

During periods of home confinement, digital devices play a crucial role in maintaining communication, accessing information, and fulfilling work and social roles. With restricted physical mobility, devices such as smartphones, tablets,

¹Link will be shared after the double-blind review.

and computers become essential for staying connected to social networks and managing day-to-day activities [31]. For instance, Ratan et al. [61] found that, during home confinement, smartphones became a vital tool for work-related communication and personal connections, often replacing face-to-face interactions.

While many studies have explored digital device usage in general contexts, the COVID-19 lockdown provided a unique opportunity to study digital usage behaviour under prolonged confinement. With workplaces and schools shifting to remote environments, the demand for digital communication tools surged [4, 18, 78]. Video conferencing platforms, social media, and messaging apps saw massive increases in usage [56, 69], and overall digital device use skyrocketed. A global survey found 70% of internet users reported increased smartphone usage during the lockdown [69], emphasizing how integral smartphones became in managing both work and personal life. Similarly, Majumdar et al. [44] observed that office workers spent significantly more time on digital devices, including smartphones, desktops, and televisions, during the lockdown. Jonnatan et al. [32] also found that mobile device usage rose during the pandemic, with urban participants reporting higher use and greater concerns about overuse, addiction, and negative health impacts.

In the post-pandemic era, many individuals continue to rely heavily on digital devices due to the persistence of remote work and online communication [28, 42]. Studies suggest that behaviours such as prolonged screen time and device overuse, which emerged during the pandemic, have continued into the post-lockdown phase [45]. This highlights the importance of understanding digital device usage patterns in today's technology-reliant environments, particularly in contexts of home confinement.

2.3 Inferring Mental Well-Being Using Mobile Sensing Technologies

Advances in mobile sensing technologies have opened new avenues for understanding human mental well-being, including mood [47, 48, 87], engagement [1, 23, 27], and mental health issues like depression and anxiety [14, 60, 80]. Many studies have integrated data from multiple sensors to improve prediction accuracy. For example, Gao et al. [23] used wearable and environmental sensors to predict student engagement, while Wang et al. [80] tracked depression using passive mobile sensing and wearables. Jaques et al. [29] modelled student mood using physiological signals and mobile behaviour.

While multi-source sensing enhances accuracy, there is increasing interest in using mobile sensing data alone, given its convenience and ability to capture a range of behaviours like physical activity, communication patterns, and location [24]. Studies have shown that mobile usage data can predict psychological states such as anxiety [60, 76], mood [47, 87], and emotion [74]. For instance, Zhang et al. [87] developed an app to detect compound emotions from smartphone usage, and Tlachac et al. [76] used call and text logs to predict anxiety and depression during the pandemic.

Research on app sequences has also provided insights into daily behaviours, such as social roles [2] and interruptibility [54, 55]. Anderson et al. [2] demonstrated that app sequences correlate with individuals' work and private roles, while Okoshi et al. [54, 55] used app sequences to predict activity breakpoints for intelligent notification delivery. However, most studies have focused on identifying breakpoints rather than modeling user activity or predicting contextual variables. Additionally, earlier studies often relied on basic aggregation methods, which fail to capture transitions between activities. Thus, **RQ1** also explores emotions and social roles by analyzing app usage patterns.

While mobile sensing technologies have shown potential in predicting well-being, gaps remain in understanding how these tools can be applied specifically during home confinement. Home confinement introduces unique behaviours—such as increased reliance on digital devices and altered perceptions of time—that differ from daily life. Although some studies [52, 73, 76] have examined psychological states during the pandemic, they focus on broader environments rather than the specific context of home confinement. Understanding this unique setting could provide valuable insights for Manuscript submitted to ACM



Fig. 1. The periods of six lockdowns and our data collection

other confinement scenarios, such as post-operative recovery, parental leave, or aging. Therefore, **RQ2** investigates how mobile sensing can model emotions during home confinement and identify key features for prediction.

3 STUDY OVERVIEW

We conducted a three-week field study in Melbourne, Australia, from 24 August 2020 to 13 September 2020, during the strictest Stage IV COVID-19 restrictions. This study use Melbourne's experience as a lens to explore broader home confinement scenarios, such as those due to medical conditions, recovery needs, or remote work. The following section details the background, participants, and data collected.

3.1 A Case Study of COVID-19 Lockdown in Melbourne, Australia

To prevent the community spread of COVID-19, Melbourne has been under six lockdowns since March 2020, with a total of 263 days of lockdowns, making it the city with the longest COVID-19 lockdown in the world [46, 72]. Our data collection was conducted during the second lockdown period with *Stage IV* restrictions (the strictest measures), which lasted from 9 July 2020 to 27 October 2020, with our study conducted from 24 August 2020 to 13 September 2020 (see Figure 1). Among the six lockdown periods, this was one of the longest and strictest lockdown periods, during which a state of disaster was declared on 2 August 2020, followed by *Stage IV* restrictions in metropolitan Melbourne [50].

Under the *Stage IV* restrictions in force at the time of our data collection, a daily curfew was implemented from 8 PM to 5 AM, forbidding anyone from leaving their homes except for special reasons. Residents were only allowed to shop and exercise within 5 km of their homes, and all students returned to home-based learning. Other restrictions that were previously applied only to specific districts were extended to the entire Melbourne metropolitan area. On 13 September 2020, several restrictions were eased in Melbourne, including reduced curfews and loosened rules on outdoor exercise and social interactions. During home confinement due to the lockdown, people showed a strong reliance on technology, particularly for school and communication [61, 69].

3.2 Procedures

Participants were recruited through advertising and comprised of university students, professional workers, academic staff, and health workers. In total, we recruited 32 participants (17 males and 15 females), where most participants were between 18-40 years old (mean = 25.44, median = 23.00, SD = 8.95). Of these, 27 participants provided smartphone data, while 13 contributed desktop data. Recruitment was primarily through social media platforms *Facebook* and Manuscript submitted to ACM

Category	Device	Total Records	Participants	Data Used for RQs
	Both	502,485	32	RQ1
App	Desktop	Total Records Participants 502,485 32 261,911 13 10 240,574 27 1,749 28 227 10 10 1,522 25 265 25	/	
	Smartphone	240,574	27	RQ2
	Both	1,749	28	RQ1
ESM	Desktop	Iotal Records Participants 502,485 32 top 261,911 13 tphone 240,574 27 1,749 28 ttop 227 10 tphone 1,522 25 265 25	/	
	Smartphone	1,522	25	RQ2
Ene-of-Day Survey	Both	265	25	RQ1

Table 1. Overview of the collected data records across devices

Discord, potentially leading to an over-representation of users from these platforms. However, as our study focused on psychological states and digital device usage during COVID-19 home confinement, not social media behaviour, we believe this does not significantly impact the validity of our findings.

Prior to data collection, participants were instructed to install the *Balance for Android* and *Balance for Desktop* apps[26, 79] on their smartphones and desktops. These apps, designed to record usage behaviors and enable experience sampling, were optimized for minimal computational, energy, and storage consumption. The apps used Experience Sampling Method (ESM) to prompt participants with a short questionnaire every 90 minutes during active device use between 7 am and 10 pm. To minimize participant burden, surveys were only triggered during active use. Additionally, the event-based ESM was triggered after 10 minutes of interaction, with follow-ups every 30 minutes. At day's end, participants received an end-of-day survey via email at 7:30 pm. After the data collection, each participant was compensated with a 25 AUD gift card as a token of appreciation for their time.

Privacy, ethics, and safety. Our field study was designed and conducted after carefully reviewing potential risks and benefits, in compliance with the city's COVID-19 safety regulations. The study has been approved by the Human Research Ethics Committee at the authors' university. To ensure privacy, all personal identifying information was anonymized before data analysis. Participants' consent was obtained prior to data collection, with a clear explanation of the study's purpose and the nature of the data being collected. We also implemented strong data security measures, including encryption of sensitive data, to further protect participants' privacy. Participants were informed of their right to withdraw at any time without penalty, and their data would be excluded from the study upon request.

3.3 Collected Data

The data collection period spanned 21 days. From the 32 recruited participants, we collected 502,485 recordings of smartphone/desktop usage, 1,749 ESM responses, and 265 end-of-day surveys. Table 1 provides an overview of the collected data across devices. To address **RQ1**, we used both desktop and smartphone data to explore how well-being was during COVID-19 home confinement. For **RQ2**, due to the limited number of participants contributing desktop data, we focused exclusively on smartphone data and its corresponding ESM responses.

Smartphone/desktop usage data. During the data collection, background services of smartphones and desktops continuously tracked the application usage (via *Android Accessibility Services*²), notifications (via *Android Notification Listeners*³), as well as the phone's states (only for smartphone, i.e. screen/battery status, ringer modes, location updates and physical activities). Specifically, smartphone app usage was recorded each time an app was opened, while desktop

²Accessibility Services: https://developer.android.com/reference/android/accessibilityService/AccessibilityService

 $^{^3}$ Notification Listeners: https://developer.android.com/reference/android/service/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/NotificationListenerService/notification/Notification/NotificationListenerService/notification/Notificat

Manuscript submitted to ACM

Category	Items	Questions / Descriptions	Options
Davias	Application usage	Foreground application name, package identifier, category	N/A
Device	Notifications	Notification arrival time, content (hashed), sender (hashed), relationship	N/A
Usuge	Phone's states	Screen status, battery, ringer modes, GPS location, physical activities	N/A
	Valence	How did you feel in the last hour? (Unhappy vs Happy)	5-Likert scale
	Arousal	How did you feel in the last hour? (Calm vs Excited)	5-Likert scale
ESM	Social role	In the last 15 minutes, I was engaged in:	Work/private/both
	Interruptibility	I am currently available for other matters:	Work/private/both/none
	Task	The task I managed to progress the most in the last 90 minutes:	Multiple choices
-	Productivity	I feel my day was productive.	5-Likert scale
	Interruptibility	I was continually interrupted today.	5-Likert scale
End-of-Day	Valence	Has the lockdown influenced your happiness today?	5-Likert scale
	Arousal	Has the lockdown influenced your energy today?	5-Likert scale
	Work-life balance	Has the lockdown impacted your work-life balance today?	5-Likert scale
Survey	Valence	How has the lockdown influenced your happiness today?	Text (optional)
	Arousal	How has the lockdown influenced your energy today?	Text (optional)
	Work-life balance	How has the lockdown influenced your work life balance today?	Text (optional)
	Difference	How different do you feel today was from yesterday?	5-Likert scale
	Difference	How was today different from yesterday?	Text (optional)

Table 2. Detailed overview of data categories, items, and response options

usage was tracked by polling the active application every second. For notifications, data captured included arrival time, hashed content (for notification length), sender (private, work, both, or none), and the participant's relationship to the sender (family, friend, work, or none). Table 2 displays an overview of the collected data for digital device usage, ESM and end-of-day survey.

ESM responses. There are some popular questionnaires to measure well-being [8, 33], and the ESM questionnaire adapted from the *Self-Assessment Manikin* (SAM) [8] was employed in this research. In the ESM questionnaire, we collected information on valence, arousal, social role, interruptibility, and tasks from users (see Table 2). Specifically, the measurement of items *valence* and *arousal* were adapted from SAM ⁴ and rated with a 5-point Likert scale, from *unhappy* to *happy* for valence and from *calm* to *excited* for arousal. These measures were used to represent the users' emotions based on the circumplex model [66]. The social role, interruptibility, and task questions were sourced from Anderson's assessment framework [3].

End-of-day survey. As shown in Figure 2(b), the end-of-day surveys were optional and included ten items designed to understand participants' well-being during the lockdown, focusing on aspects such as productivity, valence, energy, and work-life balance. Additionally, we surveyed participants on how they felt today compared to yesterday. This information helps describe the dynamics of people's experiences during the COVID-19 lockdown. The end-of-day survey questions were tailored to capture the unique context of our study and were not sourced from existing questionnaires. Each item in the end-of-day survey was rated using a 5-point Likert scale or answered with free text input (see Table 2). Specifically, items 1 and 2 were rated from 1 to 5, indicating *strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree,* and *strongly disagree.* Items 3, 4, 5, and 9 were rated from 1 to 5, indicating *a great deal, a lot, a moderate amount, a little,* and *not at all.*

4 UNDERSTANDING EMOTIONS, SOCIAL ROLES, AND WORK-LIFE-BALANCE DURING COVID-19 HOME CONFINEMENT (RQ1)

In this section, we address **RQ1** "How are emotions and social roles during COVID-19 home confinement perceived through self-reports and inferred from app usage patterns?". First, we analyse 1522 ESM responses collected via smartphones

⁴Compared to the original SAM questionnaire, we added a time constraint, asking participants to describe their feelings over the past hour. Manuscript submitted to ACM

Gao et al.



Fig. 2. The screenshots for the ESM and End-of-day survey

from 25 participants. Next, we explore the impact of the COVID-19 lockdown by examining the end-of-day survey. Additionally, we explore users' app usage behaviours through unsupervised activity recognition, identifying common activity clusters and their relationship to the perceived psychological metrics.

4.1 The Perceived Valence, Arousal and Social Roles During Covid-19 Home Confinement

Figure 3 illustrates the distribution of perceived valence, arousal, and social roles. On average, participants reported a valence of 3.46 (SD = 1.01) and arousal of 2.72 (SD = 1.11), indicating a generally calm and relaxed emotional state. This likely reflects feelings of safety and relief from the external stressors associated with the pandemic, such as the risks of infection. The absence of daily stressors—like commuting, work pressure, and social obligations—may also contribute to this sense of well-being. Furthermore, the solidarity and support networks that emerged during the pandemic likely provided additional emotional comfort. However, it is important to note that our data collection occurred during a three-week period as the COVID-19 Stage IV restrictions were lifting.Had the confinement continued longer, participants likely would have reported more negative emotions, such as boredom or frustration, as they adjusted to the new normal.

Next, we investigate the role of social roles in shaping emotional experiences during home confinement. A *social role* refers to the expectations, norms, and behaviours associated with particular social contexts [62]. These roles usually influence users' behaviours and the type of information they prefer to receive [3]. As shown in Figure 3(b), participants predominantly engaged in private-related roles, suggesting individuals tend to prioritize their personal lives and relationships during home confinement.

We further explore how emotions vary across different social roles. Figure 3(c) shows that participants reported the highest valence and lowest arousal when engaged in *private*-related roles, compared to *work*-related or hybrid roles. This suggests that personal activities, which are often less stressful and more enjoyable, foster a greater sense of well-being. Conversely, *work*-related and *hybrid* roles were associated with higher levels of stress and pressure. Welch's t-tests [88] confirm these findings, showing significant differences in valence between *private*-related and the other two roles (i.e., *private* vs. *both*, *private* vs. *work*), with p-value < 0.05. The findings offer insights into how individuals can Manuscript submitted to ACM



Fig. 3. The distribution of perceived valence, arousal, social roles and their relationship



Fig. 4. The distribution of productivity, interruption, and influence during the lockdown

structure their daily routines to boost positive emotions during home confinement. Interventions aimed at improving well-being in such situations may benefit from considering the social roles individuals are engaging in [25].

4.2 The Influence of the Home Confinement on Emotions, Social Roles and Work-Life Balance

End-of-day survey responses reveal that the lockdown had a noticeable impact on participants' valence, arousal, and work-life balance. As shown in Figure 4, 54.3%, 51.3%, and 54.7% of responses indicated that the lockdown influenced valence, energy, and work-life balance, respectively.

In terms of valence, 13 participants reported negative effects, describing feelings of anxiety, sadness, and boredom. Common complaints included missing social interactions (9 responses), the inability to go outside (10 responses), and a sense of boredom (6 responses). For instance, "I thought I might be able to see friends, but I did not", "I want to go outside, to the park", "I'm missing my family", "It's hard to say how it's specifically changed my happiness today because it's the build-up of it has gone on for so long. I'm bored and sick of my housemates and it feels like time has stopped". On the other hand, two participants reported positive effects, such as feeling happier or using the extra time for physical activity and seamless work transitions.

Upper Category	Subcategory	Тор Арр	Second App	Third App	
Business	Business	Teams	LinkedIn	Zoom	
Communication Communication		Chrome	Discord	WeChat	
	Action	Among Us	Among Us		
	Casual	Scrap Collector	IdleCourier		
	Comics	TachiyomiJ2K			
	Travel & Local	Maps	PTV		
	Video Players & Editors	YouTube	Bilibili	YouTube Vanced	
Entortainmont	Entertainment	Wow	Steam	RainbowSix	
Entertainment	Food & Drink	Uber Eats	Menulog	Mymacca's	
	Health & Fitness	Samsung Health	Nike Training		
	Photography	Photos	Gallery	Photos	
	Music & Audio	Google Play Music	Spotify	YouTube Music	
	Role Playing	Fate/GO			
	Shopping	BIGAU	Gumtree	Depop	
D. L. H. H.	Productivity	Excel	Outlook	PowerPoint	
Froductivity	Education	Canvas Student			
Reading	Books & Reference	Audible	AniDroid	MendeleyDesktop	
	News & Magazines	Joey	Twitter	Sync Dev	
Social	Social	Facebook	Instagram	Snapchat	
Tools	Tools	Gboard	Explorer	Samsung Keyboard	
Others	Auto & Vehicles	Android Auto			
	Finance	CommBank	CommSec	Up	
	House & Home	Domain	Realestate		
	Lifestyle	Samsung Pay	Tinder	Hue	
	Medical	MyTherapy	E4 realtime		
	Parenting	FamilyAlbum			
	Personalization	One UI Home	OnePlus Launcher	TouchWiz home	

Table 3. App categories (referred to the Google Play Store categories) and top three frequently used apps.

Regarding arousal, 11 participants indicated a negative impact, expressing feelings of fatigue (11 responses), lack of motivation (3 responses), or laziness (5 responses), e.g., *"I don't feel motivated to do productive things"*, *"Emotionally tired, perhaps as a result of the lockdown"*. However, two participants noted that they felt more energized during the lockdown.

In terms of work-life balance, the majority of participants reported they faced increased interruptions or distractions (14 responses), they had difficulty separating work from personal life (11 responses), and they spent more time working due to these disruptions (15 responses), with comments such as *"I worked a bit more than I should've"*, *"Everything blends together, separating work and life is really difficult"*. These findings suggest that the lockdown primarily had a negative impact on participants' work-life balance, with many struggling to manage personal and professional responsibilities. While some respondents reported positive effects, such as spending more time with family or gardening, the overall trend indicates that the lockdown negatively influenced emotional well-being and work-life balance.

4.3 Understanding App Usage Behaviours During COVID-19 Home Confinement

4.3.1 App Categorization. Apps were categorized based on historical usage logs. For Android apps, we used the categories from the Google Play Store, excluding uncategorized apps. Desktop applications were manually assigned to equivalent categories from the Google Play Store. Apps with less than one hour of cumulative usage across all participants were excluded. In total, we identified 28 app categories, which were consolidated into 8 high-level categories. A summary of the top three most frequently used apps in each category is provided in Table 3. Notably, Chrome was categorized under *Communication*, as the Google Play Store has few dedicated browser apps. Categorizing desktop applications posed a challenge, as many are used for work or software development (e.g., *PyCharm, WinSCP*). These were grouped under the *Tools* category, along with utility apps like *Explorer*.



Fig. 5. An illustration of unsupervised activity recognition using app usage data. The Segment Vector indicates the app subcategory.

4.3.2 Unsupervised Activity Recognition From App Sequences. To identify distinct user activities from app usage, we propose an unsupervised approach that combines segmentation and clustering. First, we segment the app usage sequences to capture shifts in activity, then cluster these segments to identify common patterns of behaviour.

Segmentation. To investigate the relationship between app usage and various factors, we propose *App-based Information Gain Temporal Segmentation* (AIGTS), a method that focuses on identifying different activities within app usage. Figure 5 illustrates how activities are extracted from app usage data. AIGTS is inspired by *Information Gain Temporal Segmentation* (IGTS) [67], an unsupervised method for segmenting multivariate time series. IGTS identifies splits that minimize the weighted entropy across the multivariate time series, where each variable corresponds to whether a specific app category was used.

Unlike IGTS, which focuses solely on the time series, AIGTS detects activity transitions, such as switching from Email or Browser usage (both in the *Communication* category) to browsing Reddit (in the *News and Media* category). This method can segment app usage data at various granularities, ranging from half a day to just 30 seconds. The level of segmentation is determined by the variable k, which represents the number of splits made in the time series. To ensure consistency in segment granularity across days, we compute the optimal k using the following steps: (1) Identify the knee point of the activity data for each day; (2) Calculate the average number of segments per minute for all days based on the knee point; (3) Compute the median segments per minute across all users; (4) Multiply this median by the number of minutes each user spent on their device per day and select this value as k. These steps ensure that the average segment length remains consistent across all days.

Clustering. Next, AIGTS is applied to the all-time series, resulting in a segmented list for each user. To extract features from these segments, *K-means* clustering [40] was applied. The distribution of different categories appearing in a segment was treated as an *n* dimensional vector, where *n* is the number of app categories. To identify the optimal number of segments, *Silhouette Analysis* [41] was utilized with a maximum of 20 clusters. Especially, the label indicating which cluster a segment belongs to was used as a feature representing the activity within that segment. These segment labels were derived manually from the average distribution of apps in that cluster. In this paper, a segment that belongs to a particular cluster will be referred to as an 'activity'.

Overview of User Activities. Following segmentation and clustering, we identified 16 distinct clusters representing user activities, with an average activity duration of 18 minutes. Figure 6 presents a heatmap showing the proportion of app subcategories in each cluster. The *x*-axis denotes the app categories (categories that infrequently appear in each Manuscript submitted to ACM

Gao et al.



Fig. 6. Proportion of app subcategory usage across user activity clusters. Clusters were manually named to describe the distribution of average activities within each cluster.

segment are omitted), while the *y*-axis represents the clusters (manually named such as *Deep Work*, *Hybrid Work*). The color intensity indicates the relative contribution of each app category to the cluster's overall activity distribution.

The *Device Idle* cluster, corresponding to *Personalization* apps (e.g., launchers, as shown in Table 3), represents periods when users remain on their device's home screen. We also found that the clusters *Communication Only* and *Device Idle* were the most frequent user activities during the COVID-19 home confinement, accounting for 17% and 14.6% of all activities, respectively. These activities are characterized by high proportions of communication-related apps and personalization apps, suggesting users spent considerable time on their devices in passive states or communicating with others. In addition, clusters such as *Work-Related Communication* and *Hybrid Work* exhibited higher proportions of productivity and business apps. These clusters likely reflect professional tasks and remote work activities, which were common during the lockdown period. The differentiation between *Light Work* and *Deep Work* clusters highlights variations in the intensity and type of work-related app usage.

4.3.3 Correlation Analysis of User Activity/App Usage and Well-Being. . We further investigate how user activities and app usage behaviours correlate with four key dimensions of well-being: *valence, arousal, productivity,* and *interruptions*. To assess these correlations, we performed *Spearman Rank Correlation* [17] tests between the time spent on different activities/apps and the scores for each dimension. Specifically, *valence* and *arousal* were examined for theirr relationship to the time spent on user activities in the preceding hour, while productivity and interruptions were calculated based on the end-of-day surveys.

Interestingly, as shown in Table 4, we found mostly weak correlations, with a few activities exhibiting moderate relationships. For instance, activities related to Tools Apps (e.g., phone home screens and keyboards) showed a moderate positive correlation with valence, especially among older adults. This suggests that certain habitual or passive tasks, such as interacting with productivity tools, may positively influence mood, possibly linked to work-related or routine activities. Additionally, *Surfing* activity (e.g., browsing news, including COVID-19 updates) showed a negative correlation with valence among women and the younger age group, indicating that news consumption during this period may be linked to lower mood. The correlation with productivity was notably negative for *Communication Apps* (e.g., keyboards, chat apps), but positively correlated with *Work-Related Communication* activity (e.g., business apps like Microsoft Manuscript submitted to ACM

Groups	Variable	Lowest correlation ρ	Highest correlation ρ
All	Valence	Hybrid Work (-0.07)	Tools Apps (0.11)
	Arousal	Business Tasks (-0.05)	Social Apps (0.15)
	Productivity	Communication Apps (-0.24)	Work-Related Communication (0.18)
	Interruptions	Personal Tasks (-0.15)	Communication Apps (0.29)
	Valence	Communication Only(-0.08)	Work-Related Communication (0.10)
Male	Arousal	Device Idle (-0.12)	Communication Only (0.11)
	Productivity	Communication Apps (-0.15)	Work-Related Communication (0.34)
	Interruptions	Video Streaming (-0.23)	Communication Apps (0.18)
Female	Valence	Surfing (-0.17)	Miscellaneous (0.13)
	Arousal	Hybrid Work (-0.17)	Social Apps (0.21)
	Productivity	Communication Apps (-0.30)	Personal Tasks (0.16)
	Interruptions	Personal Tasks (-0.33)	Device Idle (0.41)
Age 18-24	Valence	Surfing (-0.13)	Tools Apps (0.09)
	Arousal	Tools Apps (-0.17)	Social Apps (0.19)
	Productivity	Communication Apps (-0.21)	Hybrid Work (0.20)
	Interruptions	Social Apps (-0.10)	Communication Apps (0.16)
	Valence	Communication Only (-0.12)	Tools Apps (0.23)
A = 25	Arousal	Hybrid Work (-0.10)	Tools Apps (0.15)
Age ≥25	Productivity	Tools Apps (-0.25)	Social (0.20)
	Interruptions	Personal Tasks (-0.24)	Communication Apps (0.42)

Table 4. Top correlations between app usage/user acitivty and well-being dimensions.

Teams), highlighting that whether a communication tool is used for professional or personal purposes may influence productivity levels.

Similarly, interruptions were mostly positively correlated with *Communication Apps*, particularly among older adults. This reflects the potentially disruptive nature of non-work-related communication during the day, which may vary by age. Conversely, *Personal Tasks*(e.g., personal apps or leisure activities) were negatively correlated with interruptions, especially for older adults and women, suggesting that non-work-related tasks may help buffer against disruptions, possibly offering a sense of control or relaxation.

5 PREDICTING USER EMOTION AND SOCIAL ROLES FROM MOBILE USAGE DATA (RQ2)

In this section, we address **RQ2** "Can mobile usage data be used to model emotional well-being during home confinement, and which features are most predictive?". First, we introduce the features extracted from mobile usage data and the prediction pipeline. We then list and discuss the prediction results of the general models and individual models. Finally, we examine the significant features for predicting human well-being.

5.1 Feature Extraction

Table 5 provides an overview of the features utilized in our research. Similar to previous mobile sensing studies [39, 68, 85], we extract contextual information, notifications-related information, physical activity, app, and activities inferred from apps. To process the app-usage data, we divide it into segments based on the ESM recording frequency (set at 90 minutes) and stored it in temporal sequences. Subsequently, we extract the features from these 90-minute segments. In order to handle cyclical features such as days (e.g., Monday) and times (e.g., Morning), we transform them into sine and cosine values. Additionally, it is worth noting that app usage patterns exhibit variations between weekdays and weekends [34, 39], prompting us to conduct separate analyses for these two categories. Regarding the Manuscript submitted to ACM

Category	Feature	Description		
	weekday*	The day of the week (e.g., Monday, Tuesday)		
	time period	The time of the day (morning, afternoon, evening and midnight)		
Contextual Information	sin_weekday**,+	Sine Transformation of weekday		
	cos_weekday	Cosine Transformation of weekday		
	sin_time**,++	Sine Transformation of time period		
	cos_time**,++	Cosine Transformation of time period		
	is_weekend	Binary values describing whether is is weekend or not		
	num_noti**,++	Number of notifications		
	num_noti_interacted**,++	Number of notifications interacted		
	noti_cat ⁺⁺	Categorical Transformation of the number of notifications		
Notification	noti_intr_cat	Categorical Transformation of the number of notifications interacted		
	noti_interacted_pcg**,++	The percentage of notification interacted to the total number of notifications		
	num_communication	Number of notifications from the Communication Application		
	main_contact**	The main social relationship that contacted with		
Contextual Information sin_weekday Sin_ime*x,+i cos_weekday sin_ime*x,+i cos_weekday num_noti_interacted_pcg noti_interacted_pcg num_contain_contact** nunique_activity Physical nunique_activity**,+i num_activity**,+i nunique_app num_activity**,+i nunique_app nunique_app_cat** app_noti**,+i nunique_app_noti**,+i nunique_app*,+i+ second_cat_app**,+i second_cat_app**,+i nunique_ser_activity munique_activity* Activity Inferred nun_user_activity from Apps sec user activity	nunique_activity	Number of unique physical activities		
	main_activity**,++	The main physical activity of the users		
	num_activity*,+	Number of physical activities		
	nunique_app	The number of unique apps		
	num_app**,++	The total number of apps		
Арр	nunique_app_cat**	The number of unique app categories		
	app_noti**,++	The number of notification generated by non-communication apps		
	nunique_app_noti**,++	The number of unique apps sent notifications		
	main_cat_app ^{*,++}	The main category of apps		
	second_cat_app**,++	The second main category of apps		
Activity Inferred from Apps	nunique_user_activity	The number of unique user activities		
	num_user_activity	The number of user activities		
	main_user_activity*	The main user activity		
	sec_user_activity	The second user activity		
	entropy_user_activity	The entropy of different categories of user activity		

Table 5. The computed features from smartphone usage behaviours. For the valence model, the features selected by F-statistic $\geq 5^*$ or $\geq 10^{**}$. For the arousal model, the features selected by F-statistic $\geq 5^+$ or $\geq 10^{++}$.

time period per day, we divide it into six 4-hour segments, i.e., *Early Morning, Morning, Afternoon, Evening, Night*, and *Late Night*. Certain frequency-based features, such as notification-related or physical-activity-related features, are converted using quartiles, specifically the minimum, 25th percentile, 50th percentile, and 75th percentile. For instance, the number of interacted notifications is categorized into groups as '0', '1-3', '4-9', and '10+'. Furthermore, the physical activities considered in our study only include active physical activities like bicycling, walking, and running.

To capture user activities related to app usage, we employed our proposed method, AIGTS, as detailed in Section 4.3.2. AIGTS allows us to extract both activities and their corresponding durations. Based on the extracted information, we compute the feature 'entropy_user_activity' by determining the proportion of each activity's time portion within the entire 90-minute time interval. In addition to extracting individual features such as 'num_noti', we also generate composite features by combining two existing features. For example, the feature 'num_communication' is calculated by selecting and combining the notification sources. This composite feature counts the number of notifications originating the *Business, Communication*, or *Social* app categories. Notifications that do not belong to these app categories are counted as 'app_noti'.

5.2 Prediction Pipeline

While emotion prediction can be addressed as a classification problem, with emotion levels categorized into two or three classes based on predefined thresholds [57, 87]. We believe that utilizing regression is a better approach because valence and arousal labels are inherently ordinal data, where the relative order holds significance. In this research, we establish a regression-based pipeline to forecast the emotion scores of participants, as detailed below. Manuscript submitted to ACM

Emotion. The ESM questionnaires, completed by the participants every 90 minutes, serve as the ground truth in the regression models developed in this study. Specifically, the item 'valence' was rated on a 5-point Likert scale from 'unhappy' to 'happy', where 1 indicates 'unhappy' and 5 refers to 'happy'. Similarly, another item 'arousal' was also rated with a 5-point Likert scale from 1-5, with 1 indicating 'calm' and 5 meaning 'excited'. To predict the levels of valence and arousal, the app activities within 90 minutes prior to completing the ESM questionnaire were leveraged as features.

Regressors. In this project, we aim to develop regression models that accurately predict human emotions during home confinement. To accomplish this, we employ three commonly used regression models: *K-Nearest Neighbor* (KNN) [71], *Random Forest* (RF) [70], and *Gradient Boosting* (GB) [20]. KNN is a widely utilized model known for its versatility across various domains. It works by finding the K nearest neighbours to a given data point and using their values to predict the target variable. RF and GB, on the other hand, are ensemble models based on decision trees. These models combine predictions from multiple trees. However, there is a distinction in their approach. GB constructs trees or aggregates results in a way that minimizes the loss function, whereas RF constructs individual trees or aggregates results simultaneously.

To explore predictive capacity more comprehensively, we first develop general models using data from all participants combined. Next, recognizing that participants with similar routines or lifestyles may exhibit comparable patterns, we construct group models by dividing participants based on their self-reported habits. Finally, we analyze individual models to account for the unique characteristics and patterns exhibited by different individuals.

Feature Selection. Various methods, such as SHAP [43], LIME [64], and the *F*-statistic (*F*) [59], can be used for feature selection. In this study, we employ the *F*-statistic to determine the appropriate number of features. While most of our features are numeric, the *F*-statistic measures the univariate linear dependence between each feature and the target variable, taking into account statistical significance. Generally, a higher *F* value indicates greater importance of the feature in predicting the target variable. In our experiments, we compare regression models using three feature sets: all features, features selected based on an *F* value greater than or equal to 5, and greater than or equal to 10.

Validation. To address the limited richness of our data, we implemented *Nested Cross-Validation* (CV) [49] to train and evaluate our prediction models. Nested CV comprises an outer and an inner loop, both employing *k-fold cross-validation* [82] with different values of k. The outer loop splits the data into k folds, using each fold for testing. The inner loop further divides the remaining data into k folds, with each fold serving as a validation set to train the model and optimize hyperparameters. The best parameter settings are selected based on the inner loop results. This approach provides an accurate estimate of the model's unbiased generalization performance.

For the general models, we use a 3-fold CV in the inner loop and a 5-fold CV in the outer loop, repeated over 3 trials. The group models employ 2-fold CV in the inner loop and 3-fold CV in the outer loop. Given the limited data available for individual participants, the individual models use a simpler 5-fold non-nested CV.

Baselines and Metrics. Similar to previous human-centered studies [23, 81], we compare the proposed models with two baselines, *Mean*, which calculates the average score of user responses, and *Median*, which calculates the median value based on the data distribution and treats it as a predicted value. To evaluate the performance of regression models, we employ two commonly used metrics: *Mean Absolute Error* (MAE) and *Root Mean Square Error* (RMSE) [12]. Generally, lower values of MAE or RMSE indicate better prediction performance.

Table 6. Prediction performance of the general models for valence and arousal with different feature sets. The regression models are K-Nearest Neighbor (KNN), Random Forest (RF) and Gradient Boosting (GB), compared with Mean and Median baseline. The evaluation metrics are RMSE and MAE (with standard deviation).

Prediction Result (RMSE(SD)/MAE(SD))					
	Mean Baseline	Median Baseline	KNN Regressor	RF Regressor	GB Regressor
All features					
Valence (29)	$0.986(\pm .002)/0.824(\pm .002)$	$1.091(\pm .003)/0.767(\pm .004)$	$0.963(\pm .026)/0.776(\pm .022)$	0.903(±.025)/0.717(±.021)	$0.925(\pm .020)/0.743(\pm .012)$
Arousal (29)	$1.113(\pm.004)/0.884(\pm.004)$	$1.146(\pm .004)/0.784(\pm .004)$	$1.088(\pm .023)/0.844(\pm .019)$	$1.005(\pm .028)/0.766(\pm .020)$	$1.084(\pm .044)/0.756(\pm .037)$
Features with F-statistic >= 5					
Valence (17)	$0.986(\pm .002)/0.824(\pm .002)$	$1.091(\pm .003)/0.767(\pm .004)$	$0.956(\pm .020)/0.769(\pm .018)$	0.904(±.024)/0.717(±.021)	$0.986(\pm .020)/0.725(\pm .021)$
Arousal (14)	$1.113(\pm.004)/0.884(\pm.004)$	$1.146(\pm .004)/0.784(\pm .004)$	$1.064(\pm.027)/0.818(\pm.026)$	$1.015(\pm .024)/0.778(\pm .018)$	$1.087(\pm .032)/0.747(\pm .022)$
Features with F -statistic >= 10					
Valence (13)	$0.986(\pm .002)/0.824(\pm .002)$	$1.091(\pm .003)/0.767(\pm .004)$	0.936(±.020)/0.730(±.013)	0.912(±.024)/0.725(±.018)	$0.987(\pm .046)/0.724(\pm .019)$
Arousal (12)	$1.113(\pm.004)/0.884(\pm.004)$	$1.146(\pm .004)/0.784(\pm .004)$	$1.064(\pm.027)/0.819(\pm.026)$	$1.011(\pm .023)/0.774(\pm .017)$	1.076(±.035)/0.745(±.030)

5.3 Prediction Result

5.3.1 General Model. After excluding 19 records with invalid ESM data, a total of 1456 app-usage data records were obtained at 90-minute intervals for analysis were used to build the general models.

The prediction performance for the general models is presented in Table 6. The models used in the comparison include Mean baseline, Median baseline, KNN regressor, RF regressor, GB regressor. Among these models, the RF regressor demonstrated superior performance to both baselines in terms of MAE and RMSE for predicting valance and arousal. Interestingly, the RF regressors with different feature sets achieved comparable performance, with all feature sets yielding the best results. This can be attributed to the capability of RF to assign weights to features based on their importance, where incorporating more features leads to improved prediction performance. Furthermore, it is observed that the proposed model generally exhibited better performance in predicting *valence* compared to arousal, the difference in performance may be attributed to the complexity of arousal, which is influenced by various factors such as physiological arousal and cognitive processing. These factors may not be fully captured by mobile usage behaviours alone.

Regarding the *valence* dimension, our results indicate that the RF regressor achieved the best predictive performance across three different feature sets. For the regression model using all features, the RMSE obtained by the RF algorithm was 0.903, surpassing the Median baseline by 17.23%. Furthermore, the MAE of the RF regressor using all feature sets was 0.717, outperforming the mean baseline by 12.99% and the median baseline by 6.52%. Similar results were observed for the models utilizing the top 17 features. Regarding the *arousal* dimension, our results indicate that the RF regressor utilizing all features achieved the best performance in terms of RMSE, with a value of 1.005. This outperformed the Median baseline by 12.30%. These outcomes suggest that the selected features serve as effective predictors of *arousal* for participants during home confinement.

5.3.2 Group Models. We gathered participant work start times from pre-study questionnaires and categorized participants into two distinct groups: *EARLY* Workers (678 data points from 12 participants who typically commence work between 7:00 and 10:00 AM) and *LATE* Workers (778 data points from 13 participants who begin work after 3:00 PM). For the early workers, the average self-reported *valence* is 3.2 (\pm 0.9) and *arousal* is 2.9 (\pm 1.0). For the late workers, the average *valence* is 3.7 (\pm 1.0) and *arousal* is 2.6 (\pm 1.2). Prediction models were subsequently developed for each group using all features.

As shown in Figure 7(a), when predicting *valence* for early workers, the models performed similarly to the baselines. However, for late workers, the RF regressor outperformed the Mean baseline, achieving a 13.82% reduction in RMSE and Manuscript submitted to ACM

Living Within Four Walls



Fig. 7. Prediction performance of *valence* and *arousal* models for two user groups: *EARLY* workers (typically starting work between 7:00 and 10:00 AM) and *LATE* workers (starting work after 3:00 PM). The models are evaluated with 3-fold cross-validation, where the error bar presents the standard deviation.

a 20.36% reduction in MAE. A similar pattern is observed in Figure 7(c). For *arousal* prediction, the models demonstrated limited predictive ability for early workers but performed better for late workers, with the RF regressor reducing RMSE by 17.88% and MAE by 27.74% compared to the Mean baseline. Compared to the general models, there was no improvement in *valence* prediction for early workers, whereas for late workers, the models showed slight improvements (5% in RMSE) in both *valence* and *arousal* predictions.

These results highlight the potential value of considering user groups and routines when studying emotions. Specifically, the findings suggest that early workers' emotions may be influenced by factors beyond app usage, such as external environmental cues during the day. In contrast, late workers may exhibit a stronger link between app usage behaviours and emotions. Further insights will be explored in the later section (Section 5.4) by analyzing feature importance.

5.3.3 Individual Models. The individual models for both valence and arousal were trained using all 29 features. Due to insufficient data for some participants, only those with over 50 data points were included, resulting in a total of 13 participants and 1,220 data points. The prediction results for these models, with the ID of participants indicated on the y-axis, can be seen in Figure 8. Overall, we can observe variations in the prediction performance of valence and arousal across different participants. For instance, participant 'M110AR' consistently exhibits higher MAE and RMSE on both valence and arousal than the other participants, e.g., 'H12DPB' and 'E020KG'. This variability suggests that individual differences play a significant role in determining the accuracy of the predictions, highlighting the need to consider personalized factors when developing prediction models for emotion in home confinement situations.

Additionally, it is worth noting that while the proposed *valence* models exhibited better prediction performance than baseline models for certain participants (e.g., 'M02MMP', 'L10FGB'), there were participants who did not achieve satisfactory results. The possible reasons may be two-fold: firstly, individual differences play a crucial role in determining *valence*, and it may not be straightforward to accurately predict *valence* solely based on mobile usage behaviours for some participants. This highlights the complexity and subjectivity of *valence* and the need for a more nuanced understanding of individual emotional experiences; secondly, insufficient data availability may have limited the performance of the predictive models for some participants. Having a larger and more diverse dataset would enable better training and improve the performance of the predictive models. These findings contribute to our understanding of personalized Manuscript submitted to ACM



Fig. 8. Prediction performance of individual models using all features



Fig. 9. Top 15 features based on impurity score with RF regressors of the general models.

prediction models for well-being, which could ultimately lead to more tailored interventions and support systems to enhance individuals' well-being during periods of home confinement.

5.4 Feature Importance

To interpret the key factors influencing the prediction of *valence* and *arousal* during home confinement, we examine the feature importance of the best-performing models, which in this case are the RF regressors. The RF regressors utilize impurity-based feature importance, which measures how much including a particular feature reduces impurity across the decision tree nodes. A higher score indicates a more important feature, as it leads to a greater reduction in impurity. For brevity, we focus on the top 15 features for the general model and compare them with the top 5 features for the group models.

As depicted in Figure 9, we observe that the most important features for both *valence* and *arousal* are similar, but with different weights. These features primarily include time-related, app-category, and notification features. Both general models consistently identify 'sin_time' and 'cos_time' as the top two important features, indicating that the time period of the day greatly influences participants' *valence* and *arousal*. Additionally, the day of the week ('weekday') impacts both models, although it has less importance in the *arousal* model. Manuscript submitted to ACM

Living Within Four Walls



Fig. 10. Top 5 features based on impurity score with RF regressors of group models.

Surprisingly, the feature 'is_weekend' does not appear among the top 15 features in the general models, but it appears as important in each group models. It worth mentioning that 'time_period' is the most important feature for predicting *valence* for *EARLY* workers, but not for *LATE* workers. These findings differ slightly from the observations reported in [34], which found that app usage patterns varied greatly by time and day. The discrepancy may be attributed to the unique circumstances of home confinement, where individuals may experience a loss of time perception and a blurring of the distinction between weekdays and weekends. This aligns with the findings presented in [11], which suggest that the perception of time and daily routines can be disrupted during periods of confinement.

Besides, as shown in Figure 10, for *LATE* workers, the same five features with similar importance distributions are observed for both *valence* and *arousal* predictions, all are related to app interaction. In contrast, *EARLY* workers show a more diverse feature set; while the interacted app notification is the second most important, physical activity also contribute to *valence* prediction. This difference may reflect the broader range of daytime activities available to *EARLY* workers, potentially explaining their lower prediction performance discussed in Section 5.3.2. For *LATE* workers, the reliance on app interaction highlights a potential concern regarding phone dependency.

6 DISCUSSION AND IMPLICATIONS

This study explores the impact of home confinement during the COVID-19 lockdown on individuals' well-being, social roles, and work-life balance, using unobtrusive mobile phone sensing to gain insights into participants' emotional states and behaviors. Although the study specifically focuses on the COVID-19 lockdown, we argue that its findings have broader relevance, especially in light of the growing trend of remote work and other situations of enforced isolation.

6.1 Well-being and Social Roles During Home Confinement (RQ1)

Our analysis revealed that participants generally experienced a state of calmness and relaxation, reflected in relatively high levels of valence (positive emotional state) and lower levels of arousal (indicating a lack of intense emotional stimulation). While this suggests a degree of emotional stability during the lockdown, the study also highlights the negative effects of home confinement on work-life balance. Participants reported longer work hours and the blurring of personal and professional boundaries, which resonates with the challenges faced by remote workers in post-pandemic contexts. The segmentation of app usage data also identified key behaviors, such as increased use of communication apps and idle device time, which may reflect both increased virtual socialization and available free time.

Although the findings are rooted in the context of the COVID-19 lockdown, the implications extend beyond this unique situation. Home confinement is not limited to pandemic scenarios but is increasingly relevant to other forms of isolation or reduced mobility, such as remote work, post-operative recovery, disability-related confinement, or even digital addiction. The growing trend of telework and flexible work arrangements in the post-pandemic world further Manuscript submitted to ACM underscores the importance of understanding the emotional and behavioral impacts of prolonged confinement in various forms.

6.2 Predicting Emotional States Through Mobile Usage (RQ2)

In response to RQ2, we developed regression models that predict participants' valence and arousal based on their mobile usage data. These models demonstrated strong predictive performance, emphasizing the value of mobile phones as a tool for tracking emotional well-being in real time. Our findings indicate that time-related features (e.g., the timing of mobile phone usage) were key predictors of both valence and arousal, followed by app notifications and interactions with those notifications. This temporal sensitivity suggests that personalized interventions based on users' specific routines could enhance emotional monitoring and support during home confinement.

One key finding was the variability in the performance of prediction models across different user groups. For example, users who start work later in the day relied more heavily on mobile phone interactions, while those with early start times would benefit from models that incorporate data from other activities, such as physical exercise or interactions with environmental cues. This insight is crucial in extending the applicability of our findings to various forms of isolation beyond the pandemic. Whether it's remote workers, individuals recovering from surgery, or people experiencing limited mobility due to health conditions, the need for individualized and context-aware models remains a central takeaway.

6.3 Design Implications

Based on the findings of this study, we propose several design implications for future interventions and mobile applications aimed at supporting well-being during periods of home confinement or isolation:

Personalized Predictive Models. The success of time-related features in predicting emotional states suggests that mobile apps should incorporate personalized, time-sensitive data to offer real-time emotional tracking and tailored interventions. For example, apps could adjust their notifications and suggestions based on users' routines, ensuring that interventions are timely and relevant.

Balanced Digital Interactions. Our study found that interruptions, particularly from communication apps, may exacerbate the blurring of work-life boundaries. Future apps should focus on helping users manage interruptions, offering customizable notification settings or "focus mode" features to reduce disruptions and help maintain better separation between work and personal life.

Encouraging Healthy Digital Habits. Tools that support productivity, such as task managers and calendars, were associated with higher valence. Future designs should encourage healthy digital habits by integrating features that promote structure, goal-setting, and task management. Additionally, incorporating reminders for physical activity or mental health breaks could further support users' emotional well-being during periods of isolation.

Context-Aware Interventions. Given the differences in how users experience isolation based on their routines, digital interventions should be context-aware and adaptive. For example, users with later work hours may benefit from interventions encouraging physical activity or relaxation earlier in the day, while users with earlier start times may need more focused strategies for managing digital distractions during the day.

6.4 Generalizability of the Findings

Although this study specifically focuses on the COVID-19 home confinement context, we believe the findings can extend to a range of other isolation scenarios. While the emotional triggers—such as fear of infection in a pandemic or frustration with limited mobility in recovery—may differ, the **core psychological mechanisms** and **emotional** Manuscript submitted to ACM

dynamics of isolation are broadly similar. Our study observed patterns like emotional detachment, stress, and blurred work-life boundaries, all of which are commonly experienced during prolonged isolation, regardless of the specific circumstances.

The **emotional dynamics** of isolation—such as changes in social interactions, work-related stress, and adjusting to new routines—are consistent across scenarios like remote work, post-surgical recovery, or other forms of confinement. For example, both remote workers and those recovering from surgery face challenges with work-life balance and social isolation. Our findings about app usage—especially productivity apps promoting positive emotional states—are relevant in these contexts, suggesting that digital tools can support well-being across various isolation scenarios.

Additionally, the **mobile sensing approach** used in this study, which tracks behaviors like app usage and interaction timing, can be applied beyond COVID-19. While the types of apps and behaviors may vary, the principle of using mobile devices to track emotional states in real-time is adaptable to different situations, such as remote work or recovery periods. The ability to monitor and intervene based on individual usage patterns could be valuable in a wide range of confinement contexts.

Lastly, **context-aware interventions** are crucial for emotional well-being during isolation. Our study highlights the importance of personalized, real-time interventions—whether it's through managing digital interruptions, promoting relaxation, or encouraging physical activity. These strategies can be adapted for any form of isolation, from remote work to recovery, offering broader applicability for improving mental health.

7 LIMITATIONS

This study has several limitations that must be acknowledged, which may affect the validity of our findings [77]:

Time and Geographical Constraints (Threats to Construct and External Validity). The timing and location of data collection limit the generalizability of our results. We could not compare users' well-being and digital device usage before and during the COVID-19 lockdown due to the unpredictability of lockdown policies. Although participants reported their experiences via self-reported questionnaires, the effects of different lockdown measures and their duration may vary across regions. As the study was limited to Melbourne, the findings may not fully reflect experiences in other areas with differing restrictions.

Short Duration of Data Collection (Threats to Internal Validity). The data collection period was limited to three weeks to minimize participant burden during home confinement. While this yielded a substantial dataset, a longer observation period would have provided deeper insights into how digital device usage and emotional well-being evolve over time, particularly as individuals adapt to prolonged lockdowns. Future studies should consider longer data collection windows to capture these potential changes.

Device Usage and Untracked Data (Threats to Internal and Construct Validity). A key limitation is the potential for incomplete data due to participants using untracked devices, such as tablets or additional smartphones, or even desktops while only contributing smartphone data. These unmonitored interactions could have influenced emotional states and introduced noise into the data. The absence of comprehensive tracking across all devices limits our ability to fully assess participants' total digital behavior and emotional responses, suggesting the need for more extensive device tracking in future studies.

Sample Representativeness (Threats to External and Construct Validity). The sample size (32 participants) and recruitment methods (via Facebook and Discord ads) may have resulted in a skewed population, likely favoring tech-savvy individuals, particularly information workers. Additionally, participants who felt more positively about

being at home during the lockdown may have been overrepresented. These factors limit the generalizability of the findings to a broader population. However, despite these challenges, the data collected was sufficient for analysis.

8 CONCLUSION

In this field study, we explored how human emotions, social roles, and mobile usage behaviours were affected during the COVID-19 home confinement period. This rare event offers optimal conditions to while minimizing confounding factors like nature and physical social interactions. It contributes to understanding the mental health implications of heavy reliance on technology and work-from-home arrangements. By tracking mental well-being, the study highlights how individuals manage work-life balance and adapt their schedules or arrangements (e.g., work-from-home or half-office half-home) effectively. Additionally, it provides valuable insights into psychological states and smartphone usage behaviours during confinement, informing interventions for similar scenarios, such as post-operative home rehabilitation. Understanding how people navigate home confinement and the role of technology in their lives can guide the design of smart environments and remote support programs, improving care quality. These findings also have broader implications for understanding human behaviour during crises, supporting the development of policies and interventions for future challenges.

Despite the challenges we faced in collecting data during COVID-19 lockdowns, we believe that gaining insights into human well-being and mobile usage behaviours during home confinement is crucial. This understanding is not only valuable for policymakers in implementing effective measures during future outbreaks, but also for individuals to become more self-aware of their well-being and mental health. Moving forward, our research will expand to investigate users' desktop usage behaviour. By comparing the differences in usage behaviour across different platforms, we aim to gain a deeper understanding of people's daily lives and work-life balance.

9 ACKNOWLEDGMENTS

We sincerely thank the anonymous reviewers. We state that this manuscript has no relation to our prior publications (or concurrently submitted papers).

Living Within Four Walls

REFERENCES

- Soundariya Ananthan, Nan Gao, and Flora D Salim. 2024. Understanding Physiological Responses of Students Over Different Courses. In Proceedings of the 2024 ACM International Symposium on Wearable Computers. 104–110.
- [2] Christoph Anderson, Judith S. Heinisch, Sandra Ohly, Klaus David, and Veljko Pejovic. 2019. The Impact of Private and Work-Related Smartphone Usage on Interruptibility. In Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (London, United Kingdom) (UbiComp/ISWC '19 Adjunct). Association for Computing Machinery, New York, NY, USA, 1058–1063. https://doi.org/10.1145/3341162.3344845
- [3] Christoph Anderson, Clara Heissler, Sandra Ohly, and Klaus David. 2016. Assessment of social roles for interruption management: a new concept in the field of interruptibility. In Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct. 1530–1535.
- [4] Sam A Ankenbauer and Alex Jiahong Lu. 2020. Engaging offline communities online amid COVID-19: A case study of independent theaters. In Companion Publication of the 2020 Conference on Computer Supported Cooperative Work and Social Computing. 209–213.
- [5] Adrian Bartoszek, Dariusz Walkowiak, Agnieszka Bartoszek, and Grzegorz Kardas. 2020. Mental well-being (depression, loneliness, insomnia, daily life fatigue) during COVID-19 related home-confinement—A study from Poland. International journal of environmental research and public health 17, 20 (2020), 7417.
- [6] Daniel J. Beal. 2015. ESM 2.0: State of the Art and Future Potential of Experience Sampling Methods in Organizational Research. Annual Review of Organizational Psychology and Organizational Behavior 2, 1 (April 2015), 383–407. https://doi.org/10.1146/annurev-orgpsych-032414-111335
- [7] Bruce J Biddle. 1986. Recent developments in role theory. Annual review of sociology 12, 1 (1986), 67–92.
- [8] Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry 25, 1 (March 1994), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9
- [9] Samantha K Brooks, Rebecca K Webster, Louise E Smith, Lisa Woodland, Simon Wessely, Neil Greenberg, and Gideon James Rubin. 2020. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The lancet* 395, 10227 (2020), 912–920.
- [10] Elizabeth L Cay, N Vetter, AE Philip, and Pat Dugard. 1972. Psychological status during recovery from an acute heart attack. Journal of psychosomatic research 16, 6 (1972), 425–435.
- [11] Nicola Cellini, Natale Canale, Giovanna Mioni, and Sebastiano Costa. 2020. Changes in sleep pattern, sense of time and digital media use during COVID-19 lockdown in Italy. *Journal of sleep research* 29, 4 (2020), e13074.
- [12] Tianfeng Chai and Roland R Draxler. 2014. Root mean square error (RMSE) or mean absolute error (MAE). Geoscientific model development discussions 7, 1 (2014), 1525–1534.
- [13] Jounghwa Choi. 2020. Taylor, S.(2019). The psychology of pandemics: Preparing for the next global outbreak of infectious disease. Newcastle upon Tyne, UK: Cambridge Scholars Publishing. Asian Communication Research 17, 2 (2020), 98–103.
- [14] Philip I Chow, Karl Fua, Yu Huang, Wesley Bonelli, Haoyi Xiong, Laura E Barnes, and Bethany A Teachman. 2017. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. *Journal of medical Internet research* 19, 3 (2017), e62.
- [15] Michael Daly, Angelina Sutin, and Eric Robinson. 2020. Longitudinal changes in mental health and the COVID-19 pandemic: Evidence from the UK Household Longitudinal Study. (2020).
- [16] Krista De Castella, Philippe Goldin, Hooria Jazaieri, Michal Ziv, Carol S Dweck, and James J Gross. 2013. Beliefs about emotion: Links to emotion regulation, well-being, and psychological distress. Basic and applied social psychology 35, 6 (2013), 497–505.
- [17] Joost CF De Winter, Samuel D Gosling, and Jeff Potter. 2016. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods* 21, 3 (2016), 273.
- [18] John Dunham, Konstantinos Papangelis, Samuli Laato, Nicolas LaLone, Jin Lee, and Michael Saker. 2023. The impacts of Covid-19 on players of Pokémon GO. ACM Transactions on Computer-Human Interaction 30, 4 (2023), 1–31.
- [19] Konstantinos N Fountoulakis, Maria K Apostolidou, Marina B Atsiova, Anna K Filippidou, Angeliki K Florou, Dimitra S Gousiou, Aikaterini R Katsara, Sofia N Mantzari, Marina Padouva-Markoulaki, Evangelia I Papatriantafyllou, et al. 2021. Self-reported changes in anxiety, depression and suicidality during the COVID-19 lockdown in Greece. *Journal of affective disorders* 279 (2021), 624–629.
- [20] Jerome H Friedman. 2002. Stochastic gradient boosting. Computational statistics & data analysis 38, 4 (2002), 367-378.
- [21] Martha Leticia Gaeta, Laura Gaeta, and Maria del Socorro Rodriguez. 2021. The impact of COVID-19 home confinement on Mexican university students: Emotions, coping strategies, and self-regulated learning. Frontiers in Psychology 12 (2021), 642823.
- [22] Nan Gao, Mohammad Saiedur Rahaman, Wei Shao, and Flora D Salim. 2021. Investigating the reliability of self-report data in the wild: The quest for ground truth. In Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers. 237–242.
- [23] Nan Gao, Wei Shao, Mohammad Saiedur Rahaman, and Flora D. Salim. 2020. n-Gage: Predicting in-class Emotional, Behavioural and Cognitive Engagement in the Wild. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4, 3 (2020), 1–26.
- [24] Nan Gao, Zhuolei Yu, Yue Xu, Chun Yu, Yuntao Wang, Flora D Salim, and Yuanchun Shi. 2024. Leveraging Large Language Models for Generating Mobile Sensing Strategies in Human Behavior Modeling. In Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing. 729–735.

- [25] Michelle Hadjiconstantinou, Jo Byrne, Danielle H Bodicoat, Noelle Robertson, Helen Eborall, Kamlesh Khunti, and Melanie J Davies. 2016. Do web-based interventions improve well-being in type 2 diabetes? A systematic review and meta-analysis. *Journal of medical Internet research* 18, 10 (2016), e270.
- [26] Judith S Heinisch, Nan Gao, Christoph Anderson, Shohreh Deldari, Klaus David, and Flora Salim. 2022. Investigating the Effects of Mood & Usage Behaviour on Notification Response Time. arXiv preprint arXiv:2207.03405 (2022).
- [27] Sinh Huynh, Seungmin Kim, JeongGil Ko, Rajesh Krishna Balan, and Youngki Lee. 2018. Engagemon: Multi-modal engagement sensing for mobile games. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 2, 1 (2018), 1–27.
- [28] Taru Jain, Graham Currie, and Laura Aston. 2022. COVID and working from home: Long-term impacts and psycho-social determinants. Transportation Research Part A: Policy and Practice 156 (2022), 52–68.
- [29] N. Jaques, S. Taylor, A. Azaria, A. Ghandeharioun, A. Sano, and R. Picard. 2015. Predicting students' happiness from physiology, phone, mobility, and behavioral data. In 2015 International Conference on Affective Computing and Intelligent Interaction (ACII). 222–228. https://doi.org/10.1109/ACII. 2015.7344575
- [30] H Jeong, HW Yim, YJ Song, M Ki, JA Min, J Cho, JH Chae, et al. 2016. Mental health status of people isolated due to Middle East Respiratory Syndrome. Epidemiol Health. 2016; 38: e2016048.
- [31] Hyeon Jo and Donghyuk Shin. 2024. Influence of social and psychological factors on smartphone usage during the COVID-19 pandemic. Digital Health 10 (2024), 20552076241298482.
- [32] Livia Jonnatan, Cherisse L Seaton, Kathy L Rush, Eric PH Li, and Khalad Hasan. 2022. Mobile device usage before and during the COVID-19 pandemic among rural and urban adults. International Journal of Environmental Research and Public Health 19, 14 (2022), 8231.
- [33] Stephen Joseph, P Alex Linley, Jake Harwood, Christopher Alan Lewis, and Patrick McCollam. 2004. Rapid assessment of well-being: The short depression-happiness scale (SDHS). Psychology and psychotherapy: Theory, research and practice 77, 4 (2004), 463–478.
- [34] Kleomenis Katevas, Ioannis Arapakis, and Martin Pielot. 2018. Typical phone use habits: Intense use does not predict negative well-being. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services. 1–13.
- [35] Jagdish Khubchandani, Sushil Sharma, Fern J Webb, Michael J Wiblishauser, and Sharon L Bowman. 2021. Post-lockdown depression and anxiety in the USA during the COVID-19 pandemic. Journal of Public Health 43, 2 (2021), 246–253.
- [36] William DS Killgore, Sara A Cloonan, Emily C Taylor, Daniel A Lucas, and Natalie S Dailey. 2020. Loneliness during the first half-year of COVID-19 Lockdowns. Psychiatry research 294 (2020), 113551.
- [37] Tom Kitwood and Kathleen Bredin. 1992. Towards a theory of dementia care: personhood and well-being. Ageing & Society 12, 3 (1992), 269-287.
- [38] Sing Lee, Lydia YY Chan, Annie MY Chau, Kathleen PS Kwok, and Arthur Kleinman. 2005. The experience of SARS-related stigma at Amoy Gardens. Social science & medicine 61, 9 (2005), 2038–2046.
- [39] Tong Li, Tong Xia, Huandong Wang, Zhen Tu, Sasu Tarkoma, Zhu Han, and Pan Hui. 2022. Smartphone app usage analysis: Datasets, methods, and applications. IEEE Communications Surveys & Tutorials 24, 2 (2022), 937–966.
- [40] Aristidis Likas, Nikos Vlassis, and Jakob J Verbeek. 2003. The global k-means clustering algorithm. Pattern recognition 36, 2 (2003), 451-461.
- [41] Rosa Lleti, M Cruz Ortiz, Luis A Sarabia, and M Sagrario Sánchez. 2004. Selecting variables for k-means cluster analysis by using a genetic algorithm that optimises the silhouettes. Analytica Chimica Acta 515, 1 (2004), 87–100.
- [42] Susan Lund, Anu Madgavkar, James Manyika, Sven Smit, Kweilin Ellingrud, Mary Meaney, and Olivia Robinson. 2021. The future of work after COVID-19. McKinsey Global Institute 18 (2021).
- [43] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. Advances in neural information processing systems 30 (2017).
- [44] Piya Majumdar, Ankita Biswas, and Subhashis Sahu. 2020. COVID-19 pandemic and lockdown: cause of sleep disruption, depression, somatic pain, and increased screen exposure of office workers and students of India. *Chronobiology international* 37, 8 (2020), 1191–1200.
- [45] Stijn A A Massar, Alyssa S C Ng, Chun Siong Soon, Ju Lynn Ong, Xin Yu Chua, Nicholas I Y N Chee, Tih Shih Lee, and Michael W L Chee. 2021. Reopening after lockdown: the influence of working-from-home and digital device use on sleep, physical activity, and wellbeing following COVID-19 lockdown and reopening. *Sleep* 45, 1 (10 2021). https://doi.org/10.1093/sleep/zsab250 arXiv:https://academic.oup.com/sleep/articlepdf/45/1/zsab250/42462400/zsab250_suppl_supplementary_materials.pdf zsab250.
- [46] Alexia Matheson, Claire J McGannon, Atul Malhotra, Kirsten R Palmer, Alice E Stewart, Euan M Wallace, Ben W Mol, Ryan J Hodges, and Daniel L Rolnik. 2021. Prematurity rates during the coronavirus disease 2019 (COVID-19) pandemic lockdown in Melbourne, Australia. Obstetrics and gynecology 137, 3 (2021), 405.
- [47] Lakmal Meegahapola, William Droz, Peter Kun, Amalia De Götzen, Chaitanya Nutakki, Shyam Diwakar, Salvador Ruiz Correa, Donglei Song, Hao Xu, Miriam Bidoglia, et al. 2023. Generalization and Personalization of Mobile Sensing-Based Mood Inference Models: An Analysis of College Students in Eight Countries. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 6, 4 (2023), 1–32.
- [48] Mehrab Bin Morshed, Koustuv Saha, Richard Li, Sidney K. D'Mello, Munmun De Choudhury, Gregory D. Abowd, and Thomas Plötz. 2019. Prediction of Mood Instability with Passive Sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 3, 3 (Sept. 2019). https://doi.org/10.1145/3351233 Place: New York, NY, USA Publisher: Association for Computing Machinery.
- [49] Andreas C Müller and Sarah Guido. 2016. Introduction to machine learning with Python: a guide for data scientists. " O'Reilly Media, Inc.".
- [50] Yara Murray-Atfield and Joseph Dunstan. 2020. Curfew and exercise limits for Melbourne residents, regional Victoria moved to Stage Three lockdown. https://www.abc.net.au/news/2020-08-02/victoria-coronavirus-restrictions-imposed-death-toll-cases-rise/12515914

Living Within Four Walls

- [51] Marco Narici, Giuseppe De Vito, Martino Franchi, Antonio Paoli, Tatiana Moro, Giuseppe Marcolin, Bruno Grassi, Giovanni Baldassarre, Lucrezia Zuccarelli, Gianni Biolo, et al. 2021. Impact of sedentarism due to the COVID-19 home confinement on neuromuscular, cardiovascular and metabolic health: Physiological and pathophysiological implications and recommendations for physical and nutritional countermeasures. *European journal of sport science* 21, 4 (2021), 614–635.
- [52] Subigya Nepal, Weichen Wang, Vlado Vojdanovski, Jeremy F Huckins, Alex daSilva, Meghan Meyer, and Andrew Campbell. 2022. COVID Student Study: A Year in the Life of College Students during the COVID-19 Pandemic Through the Lens of Mobile Phone Sensing. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 42, 19 pages. https://doi.org/10.1145/3491102.3502043
- [53] Mikael Nordenmark. 2004. Multiple social roles and well-being: a longitudinal test of the role stress theory and the role expansion theory. Acta Sociologica 47, 2 (2004), 115–126.
- [54] Tadashi Okoshi, Jin Nakazawa, and Hideyuki Tokuda. 2014. Attelia: Sensing User's Attention Status on Smart Phones. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct). Association for Computing Machinery, New York, NY, USA, 139–142. https://doi.org/10.1145/2638728.2638802 event-place: Seattle, Washington.
- [55] T. Okoshi, J. Ramos, H. Nozaki, J. Nakazawa, A. K. Dey, and H. Tokuda. 2015. Attelia: Reducing user's cognitive load due to interruptive notifications on smart phones. In 2015 IEEE International Conference on Pervasive Computing and Communications (PerCom). 96–104. https: //doi.org/10.1109/PERCOM.2015.7146515
- [56] Neena Pandey, Abhipsa Pal, et al. 2020. Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. International journal of information management 55 (2020), 102171.
- [57] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment classification using machine learning techniques. arXiv preprint cs/0205070 (2002).
- [58] Delroy L Paulhus. 1991. Measurement and control of response bias. (1991).
- [59] PT Pope and JT Webster. 1972. The use of an F-statistic in stepwise regression procedures. Technometrics 14, 2 (1972), 327-340.
- [60] Haroon Rashid, Sanjana Mendu, Katharine E Daniel, Miranda L Beltzer, Bethany A Teachman, Mehdi Boukhechba, and Laura E Barnes. 2020. Predicting subjective measures of social anxiety from sparsely collected mobile sensor data. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4, 3 (2020), 1–24.
- [61] Zubair Ahmed Ratan, Sojib Bin Zaman, Sheikh Mohammed Shariful Islam, and Hassan Hosseinzadeh. 2021. Smartphone overuse: A hidden crisis in COVID-19. *Health Policy and Technology* 10, 1 (2021), 21.
- [62] Steven P Reise and Niels G Waller. 2009. Item response theory and clinical measurement. Annual review of clinical psychology 5, 1 (2009), 27–48.
- [63] Diane L Reynolds, JR Garay, SL Deamond, Maura K Moran, W Gold, and R Styra. 2008. Understanding, compliance and psychological impact of the SARS quarantine experience. *Epidemiology & Infection* 136, 7 (2008), 997–1007.
- [64] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 1135–1144.
- [65] Paula Rodríguez-Fernández, Josefa González-Santos, Mirian Santamaría-Peláez, Raúl Soto-Cámara, Esteban Sánchez-González, and Jerónimo J González-Bernal. 2021. Psychological effects of home confinement and social distancing derived from COVID-19 in the general population—a systematic review. International Journal of Environmental Research and Public Health 18, 12 (2021), 6528.
- [66] James A Russell. 1980. A circumplex model of affect. Journal of personality and social psychology 39, 6 (1980), 1161.
- [67] Amin Sadri, Yongli Ren, and Flora Salim. 2017. Information gain-based metric for recognizing transitions in human activities. Pervasive and Mobile Computing (2017). https://doi.org/10.1016/j.pmcj.2017.01.003
- [68] Zhanna Sarsenbayeva, Gabriele Marini, Niels van Berkel, Chu Luo, Weiwei Jiang, Kangning Yang, Greg Wadley, Tilman Dingler, Vassilis Kostakos, and Jorge Goncalves. 2020. Does Smartphone Use Drive Our Emotions or Vice Versa? A Causal Analysis. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376163
- [69] K Sebire. 2020. The coronavirus lockdown is forcing us to view'screen time'differently. That's a good thing. The Conversation (2020).
- [70] Mark R Segal. 2004. Machine learning benchmarks and random forest regression. (2004).
- [71] Yunsheng Song, Jiye Liang, Jing Lu, and Xingwang Zhao. 2017. An efficient instance selection algorithm for k nearest neighbor regression. *Neurocomputing* 251 (2017), 26–34.
- [72] Scott Stansfield, Arsheeya Rattan, Ben W Mol, Daniel L Rolnik, and Atul Malhotra. 2022. Impact of the COVID-19 pandemic and multiple community lockdowns on total live birth rates and preterm births in Melbourne, Australia. Australian and New Zealand Journal of Obstetrics and Gynaecology (2022).
- [73] Vedant Das Swain, Jingjing Ye, Siva Karthik Ramesh, Abhirup Mondal, Gregory D Abowd, Munmun De Choudhury, et al. 2024. Leveraging Social Media to Predict COVID-19–Induced Disruptions to Mental Well-Being Among University Students: Modeling Study. JMIR Formative Research 8, 1 (2024), e52316.
- [74] Benjamin Tag, Zhanna Sarsenbayeva, Anna L Cox, Greg Wadley, Jorge Goncalves, and Vassilis Kostakos. 2022. Emotion trajectories in smartphone use: Towards recognizing emotion regulation in-the-wild. International Journal of Human-Computer Studies 166 (2022), 102872.
- [75] Dalila Talevi, Valentina Socci, Margherita Carai, Giulia Carnaghi, Serena Faleri, Edoardo Trebbi, Arianna di Bernardo, Francesco Capelli, and Francesca Pacitti. 2020. Mental health outcomes of the CoViD-19 pandemic. *Rivista di psichiatria* 55, 3 (2020), 137–144.

- [76] ML Tlachac, Ricardo Flores, Miranda Reisch, Katie Houskeeper, and Elke A Rundensteiner. 2022. DepreST-CAT: Retrospective Smartphone Call and Text Logs Collected during the COVID-19 Pandemic to Screen for Mental Illnesses. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 6, 2 (2022), 1–32.
- [77] William MK Trochim and James P Donnelly. 2001. Research methods knowledge base. Vol. 2. Atomic dog publishing Cincinnati, OH.
- [78] Aruna Tyagi, Anoop Kumar Prasad, and Dinesh Bhatia. 2021. Effects of excessive use of mobile phone technology in India on human health during COVID-19 lockdown. *Technology in society* 67 (2021), 101762.
- [79] Niels van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. Comput. Surveys 50, 6 (Dec. 2017), 93:1–93:40. https://doi.org/10.1145/3123988
- [80] Rui Wang, Weichen Wang, Alex daSilva, Jeremy F. Huckins, William M. Kelley, Todd F. Heatherton, and Andrew T. Campbell. 2018. Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2, 1 (March 2018). https://doi.org/10.1145/3191775 Place: New York, NY, USA Publisher: Association for Computing Machinery.
- [81] Weichen Wang, Gabriella M Harari, Rui Wang, Sandrine R Müller, Shayan Mirjafari, Kizito Masaba, and Andrew T Campbell. 2018. Sensing behavioral change over time: Using within-person variability features from mobile sensing to predict personality traits. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2, 3 (2018), 1–21.
- [82] Tzu-Tsung Wong and Po-Yang Yeh. 2019. Reliable accuracy estimates from k-fold cross validation. IEEE Transactions on Knowledge and Data Engineering 32, 8 (2019), 1586–1594.
- [83] World Health Organization and others. 2020. Mental health and psychosocial considerations during the COVID-19 outbreak. Technical Report. World Health Organization.
- [84] Jiaqi Xiong, Orly Lipsitz, Flora Nasri, Leanna M. W. Lui, Hartej Gill, Lee Phan, David Chen-Li, Michelle Iacobucci, Roger Ho, Amna Majeed, and Roger S. McIntyre. 2020. Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of Affective Disorders* 277 (2020), 55 – 64. https://doi.org/10.1016/j.jad.2020.08.001
- [85] Ye Xu, Mu Lin, Hong Lu, Giuseppe Cardone, Nicholas Lane, Zhenyu Chen, Andrew Campbell, and Tanzeem Choudhury. 2013. Preference, context and communities: a multi-faceted approach to predicting smartphone app usage patterns. In Proceedings of the 2013 International Symposium on Wearable Computers. 69–76.
- [86] Robert K Yin. 2009. Case study research: Design and methods. Vol. 5. sage.
- [87] 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. https://doi.org/10.1145/3161414
- [88] Donald W Zimmerman and Bruno D Zumbo. 1993. Rank transformations and the power of the Student t test and Welch t'test for non-normal populations with unequal variances. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale 47, 3 (1993), 523.