Proceedings of the Seventeenth International AAAI Conference on Web and Social Media (ICWSM 2023)

How Circadian Rhythms Extracted From Social Media Relate to Physical Activity and Sleep

Ke Zhou[1,2], Marios Constantinides[1], Daniele Quercia[1,3], Sanja Šćepanović[1]

[1] Nokia Bell Labs, UK
[2] University of Nottingham, UK
[3] King's College London, UK
ke.zhou@nokia-bell-labs.com

Abstract

Circadian rhythm has been linked to both physical and mental health at an individual level in prior research. Such a link at population level has been long hypothesized but has never been tested, largely because of lack of data. To partly fix this literature gap, we need: a dataset on population-level circadian rhythms, a dataset on population-level health conditions, and strong associations between these two partly independent sets. Recent work has shown that affect on social media data relates to population-level circadian rhythms. Building upon that work, we extracted five circadian rhythm metrics from 6M Reddit posts across 18 major cities (for which the number of residents is highly correlated with the number of users), and paired them with three ground-truth health metrics (daily number of steps, sleep quantity, and sleep quality) extracted from 233K wearable users in these cities. We found that rhythms of online activity approximated sleeping patterns rather than, what the literature previously hypothesized, alertness levels. Despite that, we found that these rhythms, when computed in two specific times of the day (i.e., late at night and early morning), were still predictive of the three ground-truth health metrics: in general, healthier cities had morning spikes on social media, night dips, and expressions of positive affect. These results suggest that circadian rhythms on social media, if taken at two specific times of the day and operationalized with literature-driven metrics, can approximate the temporal evolution of people's shared underlying biological rhythm as it relates to physical activity ($R^2 = 0.492$), sleep quantity ($R^2 = 0.765$), and sleep quality ($R^2 = 0.624$).

1 Introduction

In his 1972 book "What Time Is This Place?", the American urban planner and author Kevin Lynch explored how humans perceive time (the human, biological sense of it), and how the environment (e.g., the physical structure of a city) influences that perception (Lynch 1972). At times, he observed, the internal rhythms of the body (the 24-hour, or circadian, cycle) may not match the collective rhythm that coordinates the actions of many people (Lynch 1972).

At an individual level, the connection between circadian factors and health has been repeatedly demonstrated, including with sleep patterns and mood regulation (Norlander, Johansson, and Bood 2005; Grandin, Alloy, and Abramson 2006; Kountouris and Remoundou 2014). Across the social sciences, there has been considerable interest in the study of how people regulate their emotions. As we shall see in Section §2, over the recent years, different studies have shown that the moods of social media users are driven in part by a shared underlying biological (circadian) rhythm (Golder and Macy 2011; Moturu et al. 2011; Murnane et al. 2015).

To go from individual level to collective level, we determined a way of capturing the circadian rhythm of a city, and then examined the relationship between the city's circadian rhythm and its health metrics. Specifically, we studied the relationship between circadian factors harnessed from Reddit and health indicators derived from a fifth of a million users who wore smartwatches for a period of three years in 18 major cities. We set out to study whether metrics reflecting circadian rhythms on social media aligned with ground-truth health data obtained from wearable devices, answering the following two research questions (RQs):

- (RQ1) Do the diurnal patterns of social media activity in cities match the circadian rhythms, or the natural progression of people's biological rhythms over time?
- (RQ2) Can we use city-level circadian rhythms derived from social media activity to predict the population-level health metrics of sleep and physical activity in a city?

In so doing, we made two sets of contributions:

- We collected 6M Reddit posts from 361K users between 2014 and 2017, and wearable readings from 233K users across 18 major cities (Section §3). From the Reddit posts, we developed five metrics that capture a city's circadian rhythm (Section §4.1). From the wearable readings, we developed three health metrics capturing sleep quantity, sleep quality, and physical activity (Section §4.2).
- We investigated the relationship between social media metrics and wearable health metrics, and found that, to a great extent, social media metrics predicted wearable health readings (Section §5), with two main findings: *1)* city-level circadian rhythms can be approximated by social media activity at night and early morning; and *2) city-level circadian metrics* are predictive of sleep quantity (with an R^2 of 0.765), sleep quality ($R^2 = 0.624$) and physical activity ($R^2 = 0.492$).

Copyright © 2023, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Those strong statistical associations indicate that the two non-fully-overlapping sets independently captured specific aspects of the city's collective behaviour rather than the two samples'. The results suggest that it is possible to approximate people's shared biological rhythms from social media data. This has important implications for developing effective population-level monitoring systems and tracking city-level circadian rhythms and health in real-time. This information can be valuable for policy makers, healthcare providers, researchers, and the public to understand city-level health and make improvements to health care systems and policy.

2 Related Work

Studying Health using Wearable Data. As consumer-grade wearables are fully equipped with body sensors, it is now possible to measure people's well-being (such as physical activity or sleep patterns) at an individual level (Park et al. 2020), or at a collective level (Aiello, Quercia, and Roitmann 2018). For example, physical activity, quantified by the number of steps, can be measured through mobile applications or wearables (Shameli et al. 2017; Aiello, Quercia, and Roitmann 2018). Good sleep has always been part of the recipe for promoting good health and well-being. Chronotypes largely determine sleeping behavior and are heavily linked to genetics (Vink et al. 2001). In addition to genetics, researchers have studied an array of other factors that affect sleep, from environmental factors (Roenneberg, Wirz-Justice, and Merrow 2003) to social ones (Wittmann et al. 2006). Measurements of sleep behavior rely on devices such as actigraph or commercial wearables that estimate human rest/activity cycles. By collecting sleep variables using actigraphy, for example, Miller *et al.* (Miller et al. 2015) investigated the relationship between chronotype and positive affect and found that night chronotypes (i.e., late sleepers) have significantly lower positive affect and are more likely to suffer from depression. Mobile-based solutions have also been proposed to detect sleep phases and quantify sleep quality (Krejcar, Jirka, and Janckulik 2011; Hao, Xing, and Zhou 2013; Min et al. 2014). For example, Min *et al.* (Min et al. 2014) utilized movement, ambient light, sound, screen state, apps usage, and battery status to quantify sleep quality. Finally, leveraging large-scale wearable data, Aiello *et al.* (Aiello, Quercia, and Roitmann 2018) showed that our body's natural rhythms get disrupted by exogenous events (e.g., elections, festivities), and that was the case for both physical activity and sleep.

Monitoring Health with Social Media. Past research has looked into the relationship between social media and various health outcomes such as stress (Lin et al. 2016; Guntuku et al. 2019; Saha et al. 2019), sleep quality (insomnia) (Jamison-Powell et al. 2012; McIver et al. 2015), mental health (Dos Reis and Culotta 2015), and obesity (Karami et al. 2018). For example, linguistic features have been used to analyze stress from Twitter at scale (Guntuku et al. 2019). Guntuku *et al.* adapted machine learning and natural language techniques to construct classifiers for predicting stress levels of social media users (Guntuku et al. 2019). Social media expressions of stress and their relationship with gun violence on university campuses was investigated in the 2017 Saha and De Choudhury's study (Saha and De Choudhury 2017). The researchers found that Reddit had significant higher levels of stress expressions after gun violence events. These expressions reflected *"reduced cognition, higher self pre-occupation and death-related discussion."*

Circadian Rhythm and Health. The human body's biochemistry varies predictably throughout the day, and pulses in cycles (Carrier and Monk 2000) – we sleep or wake up, we are hungry or full, we are alert or tired. The most dominant period in a person's rhythms is the circadian cycle. This was first noted by Curt Richter, a biologist studying the relations among activity, sleep, and the 24-hour clock (Richter 1967). The circadian rhythm not only dictates when we sleep or wake up, but also significantly contributes to, for example, our dips and peaks of alertness throughout a day (Abdullah et al. 2014). Prior studies found associations between circadian rhythms and health, both physically (Cho 2001) and mentally (Biss and Hasher 2012; Kreitzman and Foster 2011; Facer-Childs et al. 2019). For example, 'morning larks' were found to be higher in positive affect than those who woke up late (Biss and Hasher 2012). Researchers reset the undesirable late timing of 'night owls' in a study and consequently observed a positive impact on cognitive performance and mental health (Facer-Childs et al. 2019). Furthermore, a positive but weak association between morning alertness and sleep quality was reported in (Reid, Maldonado, and Baker 2002), suggesting that, if people sleep well at night, they tend to have higher alertness levels the next day. Walsh et al. (1991) found that the best measure of sleep quality in healthy individuals, beyond the amount of sleep, is one's alertness level during waking hours. Persistent disruption of the biological rhythm can have serious consequences for physical and mental well-being (Foster and Wulff 2005). For example, the effects of sleep debt, similar to crossing time zones, can cause temporal lobe atrophy (amnesia) and spatial cognitive deficits (Cho 2001). Circadian rhythms were successfully derived from social media use (Golder and Macy 2011; Aiello, Quercia, and Roitmann 2018). For example, by analyzing millions of public Twitter messages (Golder and Macy 2011), Golder and Macy found that individuals' mood deteriorates as the day progresses, consistently with circadian factors. Not only these patterns hold true on Twitter, but, more recently, the same trends were observed by analyzing millions of Spotify's music plays (Park et al. 2019).

Literature Gap. While social media has been found to contain key behavioral markers that reflect a shared biological rhythm, it is yet unclear the extent to which these markers are actually tracking a city's collective health. Therefore, we set out to study whether metrics reflecting circadian rhythms on social media matched ground-truth health data obtained from wearable devices.

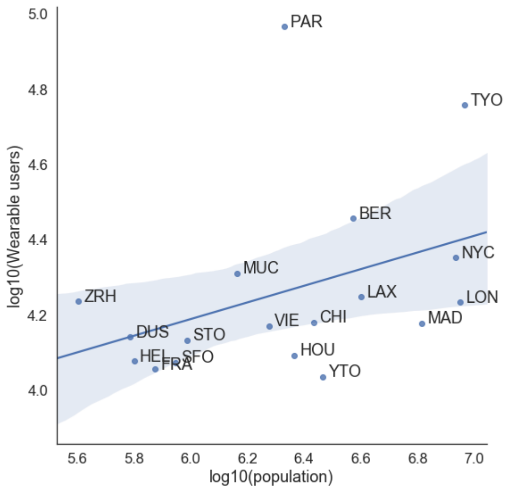


Figure 1: Relationship between the number of wearable users and city population (log-transformed): Spearman correlation $r = .46$ and $p = 0.05$.

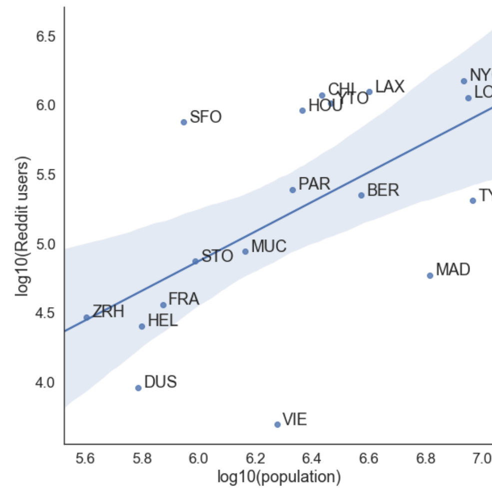


Figure 2: Relationship between the number of Reddit users and city population (log-transformed): Spearman correlation $r = .65$ and $p = 0.003$.

3 Datasets

3.1 Smartwatch Readings

We obtained three types of data from commercial smartwatches worn by 232,707 unique users (55% male with 42 years as the median age) between 2014 and 2017 across 18 major cities: Berlin (BER), Dusseldorf (DUS), Chicago (CHI), Frankfurt (FRA), Helsinki (HEL), Houston (HOU), Los Angeles (LAX), London (LON), Madrid (MAD), Munich (MUC), New York (NYC), Paris (PAR), San Francisco (SFO), Stockholm (STO), Tokyo (TYO), Vienna (VIE), Toronto (YTO), and Zurich (ZRH). We relied on two types of reading:

1. **Number of steps.** We obtained the average daily number of steps per user.
2. **Bed-in and bed-out time.** The bed-in time is the estimated time (in units of Unix timestamp) at which a user gets in bed. For example, a value of 23:00:00 means that the user went to bed at 11:00 PM, whereas a value of 25:04:00 means the user went to bed at 1:04 AM. The bed-out time is the estimated time (in units of Unix timestamp) at which a user gets out of bed. For example, a value of 07:04:00 means the user got out of bed at 7:04 AM.

To minimize any confounding effects, we ascertained that each device's software and hardware remained the same throughout the data collection period. Additionally, during the three-year data collection period, the number of steps and sleep readings were monitored almost continuously throughout the year (90%+ of the days), thus ensuring negligible missing data. We also correlated the number of wearable users with the city's population size (Figure 1). We found a Spearman's rank correlation of 0.46, suggesting a moderate representation of our wearable users in the 18 cities, not least because Paris and Tokyo were overrepresented: Paris is the headquarter of the wearable company, and Tokyo is an early adopter of wearable devices, beyond the case of this specific company[1].

Data Ethics of Smartwatch Logs. The data processing in this study is compliant with the smartwatch company's terms and conditions. Additionally, in accordance with the General Data Protection Regulation (GDPR), no researcher involved in the study could have tracked the identity of any user by any means, and all readings were obtained and analyzed at an aggregated level.

3.2 Reddit Data

Reddit is a public discussion website structured in independent communities—called subreddits—dedicated to a broad range of topics (Medvedev, Lambiotte, and Delvenne 2017). Users can upload posts to a subreddit, write threads of replies to existing posts, and upvote or downvote posts and replies (Baumgartner 2015). It is an online forum in which users generally voice their opinions or feelings. We processed the Reddit data in three steps:

- **Removing extreme and long-tail users**: As there are multiple 'moderator' or 'bot' users posting on Reddit, to avoid biasing our results, we removed (extreme) Reddit users that posted either more than 1K messages (likely bots) or only one message (long-tail users).
- **Adjusting timestamps**: To make the temporal analysis possible, we adjusted the timestamps of each user's posts to his/her local time zone.
- **Geo-referencing Reddit posts**: We used the method from Balsamo et al. (Balsamo, Bajardi, and Panisson 2019) to assign users to their geographical location. Specifically, we first identified a list of subreddits that can be matched to cities (e.g., r/london). Then, for each user

[1]https://www.wareable.com/wearable-tech/wearables-becoming-huge-in-asia-544

who posted at least once in these subreddits, we assigned the user to the corresponding city and country. Note that if a user posted in multiple countries, we assigned the user to the country having his/her majority of posts. This ensures that users who write posts in a city's subreddit because they plan to move there or visit as tourists are not assigned to that city. The heuristic for assigning users to states in the US has been found to be accurate and to approximate population size to a great extent (Balsamo, Bajardi, and Panisson 2019). In our case, we also found that the number of users assigned to a city is correlated with the city's population size, achieving a Spearman's rank correlation of 0.65. Yet, in Figure 2, we observe that San Francisco (SFO) is over-represented on Reddit (given its focus on tech), while Vienna (VIE) is under-represented. Despite that, even for under-represented cities, we ascertained that we had sufficient data for computing the five circadian metrics. We did so by reducing the number of points at a minimum. That is, for each metric, we found the minimum number of data points. This was equal to the number of data points in the least represented city. For the remaining cities, we created reduced sets by removing data points at random (these cities were all described by the same minimum number of points). We then computed the correlation between the metric on the full set and the metric on the reduced set. For all metrics, the correlations were above 0.90, suggesting that even the least represented city had a sufficient number of points to quantify the metric.

By applying these three steps, we then obtained the Reddit posts between 2014 and 2017 across the 18 cities (same cities as in Section 3.1), spanning across three continents and with high Reddit penetration rates. In total, we collected 5,548,984 posts from 361,196 users.

3.3 Official Urban Statistics

In our study, we controlled for two main variables. First, for GDP per capita, collected from the Organization for Economic Co-operation and Development (OECD) regional statistics, and log-transformed using the logarithm (log_gdp) due to its skewed distribution ($min = 11.31, max = 14.38, \mu = 12.77, \sigma = 0.94$). Second, for population size, collected from the cities' Wikipedia pages, and log-transformed using the logarithm (log_pop) due to its skewed distribution ($min = 12.91, max = 16.04, \mu = 14.62, \sigma = 0.96$).

Other socio-economic variables could have been added (e.g., presence of tech workers) but would have not been standardized to then be comparable across the 18 cities. As for our two controls, the gross domestic product (GDP) is a common socio-economic factor that reflects a region's wealth, and city population size has been found to be predictive of a wide variety of processes, including urbanization, economic development, and knowledge creation (West 2017). Indeed, patent production, personal income, and electrical cable length have all been shown to be power law functions of population size with scaling exponents that fall into distinct universality classes (Bettencourt et al. 2007).

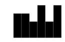
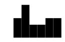
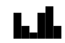
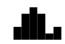
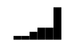
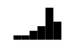
Variable	Distribution	Min	Mean	Max
log population		12.91	14.57	16.04
log GDP		11.31	12.77	14.38
morning peak		78.0	6704.6	22491.0
night dip		0.77	0.93	1.00
morning larks		0.11	0.27	0.42
night owls		0.05	0.11	0.18
positive affect		0.10	0.23	0.39
sleep quantity		6.90	7.59	8.02
sleep quality		0.50	0.65	0.70
physical activity		5915.6	7057.4	7606.1

Table 1: Frequency distributions and statistics for all metrics, including official urban statistics, social media metrics, and wearable health metrics. The "Distribution" column shows the frequency distribution of the variable within the dataset (x-axis ranges from Min to Max, and y-axis shows the frequency).

4 Methodology

We processed the previously described sets of data to compute metrics from social media (Section 4.1), and from wearable data (Section 4.2). The resulting statistics are summarized in Table 1.

4.1 Social Media Metrics

Building upon the literature (Golder and Macy 2011; Park et al. 2019), we were able to compute five metrics that were previously hypothesized to relate to collective circadian processes. These studies showed that individuals tend to start the day in a positive state of affect (such as enthusiasm, alertness, and activeness), but this deteriorates as the day progresses. Since the metric designs have been validated in similar contexts, such as Twitter (Golder and Macy 2011) and Spotify music streaming (Park et al. 2019)), we adopted their methodologies for designing our social media based metrics on Reddit.

Online Activity. Circadian rhythm is an internal clock that coordinates an individual's daily cycle. It has been found that, for a typical person, alertness has 4 cycles of rise and fall in a day (Figure 3), identifying four time frames of "peak-fall-rise-dip" (Richter 1967): a [6am-11am] morning peak of alertness, a [11am-3pm] mid-day fall, a [3pm-8pm] evening rise, and a [8pm-6am] night dip. For each timeframe, we computed a city c's temporal change in online activity based on Reddit posting activity:

$$activity\_change(c, h[i-1], h[i]) = \frac{x(c,h[i]) - x(c,h[i-1])}{x(c,h[i-1])} \quad (1)$$

where $h \in \{6,11,15,20\}$ (shown with the vertical lines in Figure 3), $x(c, h)$ is the number of Reddit posts in hour h in city c. To ensure that a few active users would not dis-

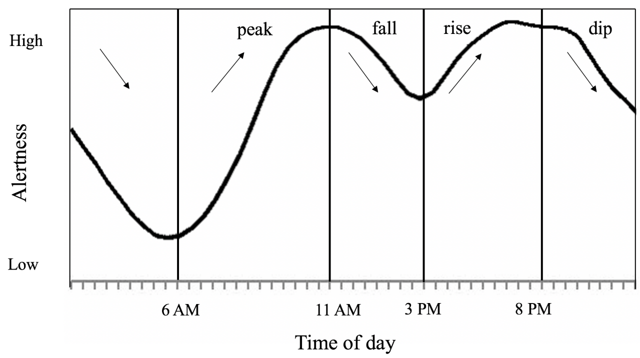


Figure 3: The circadian rhythm is an internal clock coordinating an individual's four daily cycles of alertness: a [6am-11am] morning peak of alertness, a [11am-3pm] mid-day fall, a [3pm-8pm] evening rise, and a [8pm-6am] night dip.

proportionately drive this metric, we experimentally ascertained that a sufficiently large number of users were active in each timeframe under study. Prior studies found that sleep patterns have a significant impact on one's health (Luyster et al. 2012). Going to bed early at night and waking up early in the morning has been found to have health benefits (Reid, Maldonado, and Baker 2002; Biss and Hasher 2012; Walsh et al. 1991). Therefore, we are particularly interested in the online activity change in the morning from 6am to 11am (expected to be an online activity peak) (Walsh et al. 1991), and at night from 8pm to 6am (expected to be an online activity dip) (Reid, Maldonado, and Baker 2002). Furthermore, as we shall see in Section 5.1, the city daily online activity computed on social media (Figure 5) matched the individual daily alertness (Figure 3) during the [6am-11am] morning peak of alertness, and the [8pm-6am] night dip. By contrast, we did observe neither the [11am-3pm] mid-day fall nor the [3pm-8pm] evening rise, suggesting that circadian factors extracted from social media are more likely to reflect how users regulate sleep - which is partly how the daily cycle emerges - rather than how alert they are. Given this initial experimental finding, we then focused on the morning peak and the night dip, and computed them as:

$$morning\_peak(c) = activity\_change(c, 6, 11) \quad (2)$$

$$night\_dip(c) = -activity\_change(c, 20, 6) \quad (3)$$

where *morning_peak* of online activity is high, if the online activity rises fast in the morning (e.g., in a city in which Reddit users are suddenly active in the morning); while *night_dip* is high, if online activity falls fast at night (e.g., in a city in which Reddit users are suddenly going offsite and, potentially, offline at night).

Chronotypes. Circadian rhythms vary between individuals who can, for example, be "early birds" (morning types) or "night owls" (evening types) (Golder and Macy 2011). Following Golder and Macy (Golder and Macy 2011), we allocated each user to either of the four six-hour chronotypes based on the time (s)he was most active on Reddit. Given

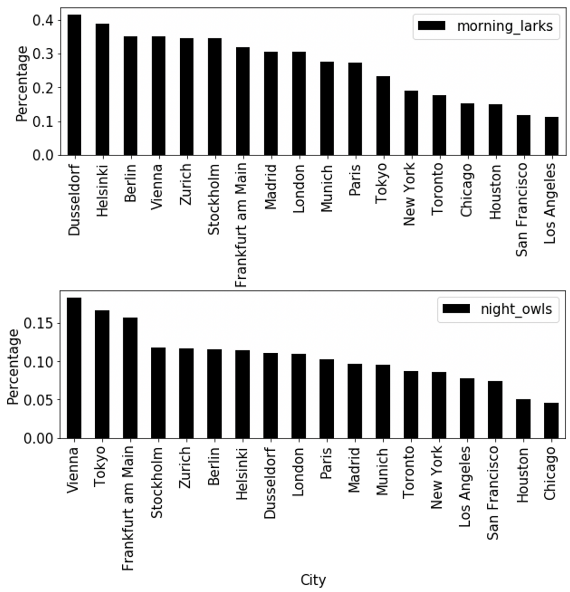


Figure 4: Percentage of users who are morning larks (top) and night owls (bottom) across the 18 cities.

our experimental focus on both night and early morning, we considered the two corresponding chronotypes of "morning larks" [6am-12pm] and "night owls" [12am-6am], which past studies at individual level associated with sleep quality (Walsh, Repa, and Garland 2021) and cognitive performance (Facer-Childs, Boiling, and Balanos 2018). We calculated their prevalence in city c as:

$$morning\_larks(c) = \frac{\sum_{u \in U(c)} \underset{p}{argmax}\ x(u,p) \cdot I(p, \text{“}morning\text{”})}{|U(c)|} \quad (4)$$

$$night\_owls(c) = \frac{\sum_{u \in U(c)} \underset{p}{argmax}\ x(u,p) \cdot I(p, \text{“}night\text{”})}{|U(c)|} \quad (5)$$

where $x(u,p)$ is the number of u's posts during period p; p is any of the four time periods $p \in \{\text{“}morning\text{”}, \text{“}afternoon\text{”}, \text{“}evening\text{”}, \text{“}night\text{”}\}$; $I(p,p')$ is an indicator function denoting whether $p = p'$ (in which case $I = 1$), or not ($I = 0$); $U(c)$ is the set of Reddit users in city c; and $|U(c)|$ is the total number of those users. To put formula (4) in plain English, we can say that its numerator sums 1's for all the users who are most active during the morning [6am-12pm], and its denominator normalizes that quantify by the total number of users in the city. Formula (5) does the same but for the users who are most active at night [12am-6am]. For our 18 cities, we observed a sufficient variability for both chronotypes (Figure 4), with, say, Vienna having considerable fractions of larks and owls, and, by contrast, with Chicago enjoying a large fraction of more "regular" sleepers.

Positive Affect. Individuals regulate mood to cope with the daily demands and to ultimately function in productive ways. Following (Golder and Macy 2011), we captured an individual's overall affect as his/her mean positive affect across all hours:

$$positive(u) = \frac{1}{||H||}\Sigma_{h \in H} PA(u,h) \quad (6)$$

where $PA(u, h)$ is the percentage of u's posts that express positive affect in hour h, and H is the set of hours in the study period. By then aggregating these individual's overall affect values for each city c, we obtained:

$$positive\_affect(c) = \frac{\sum_{u \in U(c)} positive(u)}{|U(c)|} \quad (7)$$

where $U(c)$ is the set of Reddit users in city c, and $|U(c)|$ is the total number of these users.

4.2 Wearable Health Metrics

From the three types of smartwatch readings, we developed three health metrics at city level.

Sleep Quantity. This metric is the average of the mean daily hours each user spent in bed. The mean daily hours an individual spends in bed is often used as a proxy for the quantity of rest the individual gets at night (Abdullah et al. 2016; Ancoli-Israel et al. 2003; Walch, Cochran, and Forger 2016). The higher the value of this metric, the more time the 'typical' (average) wearable user in c slept:

$$sleep\_quantity(c) = \frac{\sum_{u \in U'(c)} (bed\_out(u) - bed\_in(u))}{|U'(c)|} \quad (8)$$

where $(bed\_out(u) - bed\_in(u))$ is the daily number of hours, on average, user u spent in bed, and $|U'(c)|$ is the number of wearable users in city c. The values of bed_in and bed_out are computed by the wearables based on one's past behavior and current movements. As such, bed_in starts not when an individual simply goes to bed but when (s)he goes to bed and does not produce any sudden physical movement.

Sleep Quality. This metric captures the extent to which wearable users in a given city aligned with the eight hour golden rule of rest (Chaput et al. 2013). Unlike the sleep quantity metric, sleep quality measures the divergence of people's daily sleep time. The higher the value of this metric, the better the city's wearable users slept (i.e., the lower their collective deviation from the rule):

$$sleep\_quality(c) = 1/(\frac{\sum_{u \in U'(c)} |bed\_out(u) - bed\_in(u) - 8|}{|U'(c)|}) \quad (9)$$

where $sleep\_quality(c)$ is *inversely* proportional to a quantity whose numerator sums each user u's number of hours above/below the recommended 8 hours, and it does so over all the users in city c, while its denominator re-scales by the total number of users in c.

Physical Activity. This metric captures the amount of physical activity in terms of average daily number of steps per wearable user within a city. The higher the value of this metric, the more steps the 'typical' wearable user in city c did on a daily basis:

$$physical\_activity(c) = \frac{\sum_{u \in U'(c)} steps(u)}{|U'(c)|} \quad (10)$$

where $steps(u)$ is the average number of u's daily steps.

5 Results

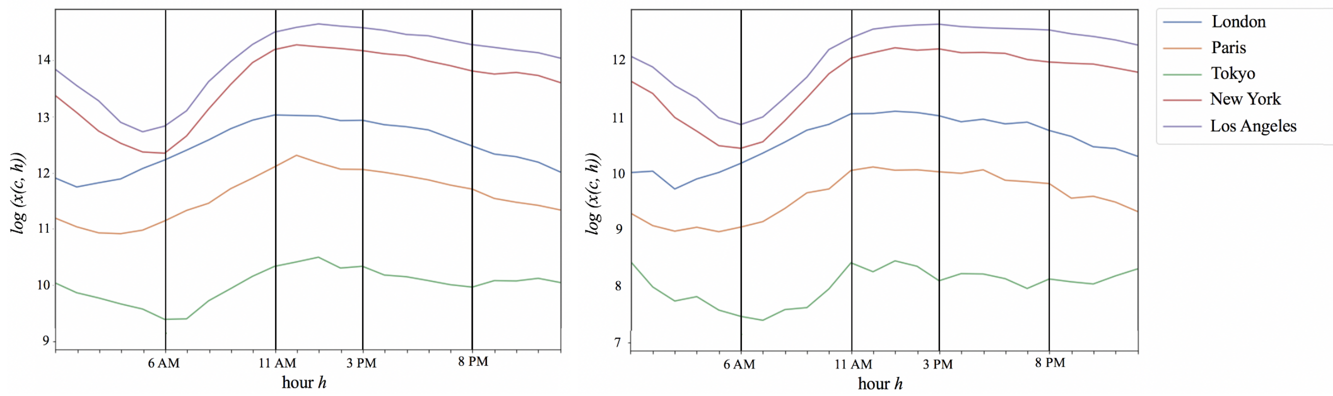
5.1 Validating Social Media Metrics

To establish the internal validity of our approach, we plotted the log-transformed number $x(c, h)$ of posts at each hour h in each city c. For presentation purposes, Figure 5 shows the diurnal pattern of five major and representative cities in our dataset (i.e., two main European cities, one east-coast US city, and one west-coast US city, and an Asian city); the remaining thirteen cities followed closely the diurnal patterns of the five. Across all cities, circadian rhythms for both night and early morning were systematically confirmed. From early morning till noon, Reddit users were considerably active (matching the hypothesized [6am,11am] morning peak). From noon until midnight, there was a steady and slight drop (not reflecting the hypothesized mid-day fall and evening rise of alertness). Finally, from midnight till 5am on weekdays (and till 6am on weekends), there was a sharp decay in activity (matching the hypothesized [8pm,6am] night dip). Overall, these results suggest that the number of daily posts at each hour h in c ($x(c, h)$) captured sleeping patterns rather than more generalized alertness patterns. We will show that the metrics computed during both night and early morning, unlike those computed on the remaining parts of the day, were indeed predictive of the three health metrics, as suggested by individual-level studies (Reid, Maldonado, and Baker 2002; Biss and Hasher 2012; Walsh et al. 1991).

Although the overall temporal pattern was highly robust, there were interesting between-group differences. Reddit users in Europe generally rose earlier than those in the US and Asia: posting activity in London and Paris started as early as 5am, while that in American cities or Asian ones started at around 7am. To compensate for that, American cities were more active at night (only experiencing a drop at as late as 1am).

5.2 Cross-correlations Between Wearable Health Metrics and Social Media Metrics

Despite having ascertained the validity of the social media metrics, we have to note that these metrics were not orthogonal to each other. On the contrary, some of them correlated with each other (Figure 6). Morning peak of online activity negatively correlated with prevalence of night owls [$r_{18}(\text{morning peak, night owl}) = -0.723$, $p < 0.01$] and, surprisingly, of morning larks [$r_{18}(\text{morning peak, morning larks}) = -0.822$, $p < 0.01$]. These two negative correlations indicated that morning peaks are created by a regular sleeping pattern: one without night owls and morning larks. In fact, by removing night



(a) weekdays (b) weekends

Figure 5: The log-transformed number $log(x(c, h))$ of posts at each hour h in each city c. For all cities, its values matched the daily evolution of individual circadian alertness (shown in Figure 3) during two times of the day: early morning (6am to 11am) and late night (8pm to 6am). As for comparison between weekdays and weekends, Reddits users generally remained active for one more hour during weekend nights, which is consistent with prior work (Park et al. 2019).

owls and morning larks from our dataset, we observed even more marked morning peaks.

The health metrics correlated with each other as well. Sleep quantity (bed duration) positively correlated with sleep quality (inverse 8-hour rule) [$r_{18}(\text{sleep quantity}, \text{sleep quality}) = 0.494$, $p < 0.05$], and moderately with physical activity [$r_{18}(\text{sleep quantity}, \text{physical activity}) = 0.125$, $p >= 0.1$].

Finally, by examining the pairwise rank correlations among all variables (including GDP and city size), we observed that: healthier cities (e.g., those with better sleep) were generally characterized by positive circadian markers (e.g., [$r_{18}(\text{night dip}, \text{sleep quantity}) = 0.653$, $p < 0.01$]); and wealthier and larger cities (higher GDP and population) by less satisfactory sleep (e.g., [$r_{18}(\text{log GDP}, \text{sleep quality}) = -0.746$, $p < 0.01$] and [$r_{18}(\text{log population}, \text{sleep quality}) = -0.622$, $p < 0.01$]) but by slightly more physical activity ([$r_{18}(\text{log population}, \text{physical activity}) = 0.209$, $p >= 0.1$]). These results are in line with the literature reporting that city life negatively impacts quality of sleep (e.g., insomnia is more prevalent (Tähkämö, Partonen, and Pesonen 2019)) but may support a wider variety of options for physical activity (Althoff et al. 2017).

5.3 Predicting Wearable Health Metrics From Social Media Metrics

To test the external validity of our approach, we developed models that predicted the three health metrics from the five social media metrics. Specifically, we developed nine multiple regression models that predict sleep quantity, sleep quality, and physical activity from the circadian rhythm metrics derived from social media, while controlling for a city's wealth and population size. For each dependent health variable, we developed a set of three models: *1)* a baseline model that used log-GDP and log-population as predictors, *2)* a full model that included all the social media metrics and the two control variables, and *3)* an economic model that included variables selected using a stepwise feature selection method. This method is called StepAIC (Zhang 2016) and selects the model with the smallest AIC (Akaike Information Criteria) by iteratively adding predictors that decrease AIC and/or removing those that increase it. In other words, it returns the model that strikes the right balance between being economical (fewer independent variables) and having predictive power. All variables in our models were scaled between 0 and 100 to ease the interpretation of the correlation coefficients.

Predicting Sleep Quantity. A baseline model to predict sleep quantity from log-GDP and log-population achieved an Adjusted R^2 of 0.226 (M1), while the full model achieved an Adjusted R^2 of 0.740 (M2). The best predictive and most economical model achieved an Adjusted R^2 of 0.765 (M3), and the StepAIC function selected morning peak of online activity, night dip of online activity, prevalence of morning larks, and GDP as the four most important predictors. The beta coefficients (of M3 in Table 2) indicate that: a 10% increase in the morning peak of online activity led to 9.3% increase in sleep quantity; a 10% increase in the night dip of online activity led to 5.7% increase in sleep quantity; and a 10% increase in prevalence of morning larks led to 5.4% increase in sleep quantity. These results suggest that, in cities in which our wearable users slept for longer, posting on Reddit generally sharply decreased at night and increased quickly in the morning. In addition, as GDP increased, wearable users generally slept less ($\beta = -0.519$), which is in line with the literature (as the two economists Biddle and Hamermesh put it, *'the more you can earn, the more worthwhile it may seem to sacrifice sleep for work'* (Biddle and Hamermesh 1990)).

Predicting Sleep Quality. A baseline model to predict sleep quality from log-GDP and log-population achieved an Ad-

Feature	sleep quantity M1 *Baseline*	M2 *Full*	M3 *StepAIC*	sleep quality M4 *Baseline*	M5 *Full*	M6 *StepAIC*	physical activity M7 *Baseline*	M8 *Full*	M9 *StepAIC*
morning peak		1.271**	0.932***		1.023*			0.076	
night dip		0.837**	0.570***		0.695**	0.193		-0.107	
morning larks		0.368	0.540***		0.376	0.547*		1.183*	
night owls		0.034			-0.059	-0.441*		0.096	
positive affect		0.392			0.642*			0.234	
log population	-0.267	0.101		0.084	0.129		0.342	-0.312	
log GDP	-0.332	0.809	-0.519**	-0.788**	-0.926	-0.371	-0.253	1.033	0.821***
intercept	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
adjusted R^2	0.226	0.740	**0.765**	0.463	**0.667**	0.624	-0.078	0.253	**0.492**

Table 2: The regression analysis predicting the three health metrics (sleep quantity, sleep quality, and physical activity) from the five social media circadian rhythm metrics, while controlling for log GDP and log population. Significant values are marked with a number of *'s based on their significance levels (i.e., *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$).

justed R^2 of 0.463 (M4), while the full model achieved an Adjusted R^2 of 0.667 (M5), which is the best performing model. Slightly underperforming the full model, the most economical model (M6) achieved an Adjusted R^2 of 0.624. The StepAIC function selected the prevalence of morning larks and night owls as the two most significant predictors. The beta coefficients (of M6 in Table 2) indicate that: a 10% increase in prevalence of morning larks led to 5.4% increase in sleep quality; and a 10% increase in prevalence of night owls led to 4.4% decrease in sleep quality. In addition, GDP and night dip of online activity also played a role. As we hypothesized, in cities with a higher prevalence of morning larks and lower prevalence of night owls, users slept better. In addition, users in cities with higher GDP generally experienced more unsatisfactory sleep, in line with our previous results for sleep quantity.

Predicting Physical Activity. A baseline model to predict physical activity from log-GDP and log-population achieved an Adjusted R^2 of -0.078 (M7), on par with a random predictor. In contrast, the full model (M8) yielded a slightly increase in the Adjusted R^2 of 0.253. The best predictive and most economical model of physical activity interestingly achieved an Adjusted R^2 of 0.492 (M9). The StepAIC function selected the prevalence of morning larks and GDP as the two most important predictors. More specifically, the beta coefficients (of M9 Table 2) indicate that: a 10% increase in prevalence of morning larks led to 11% increase in physical activity; and a 10% increase in in log-GDP led to 8.2% increase in physical activity. These results suggest that, in wealthier cities with a higher prevalence of morning larks, where better health was expected (Facer-Childs, Boiling, and Balanos 2018), we indeed found that our wearable users engaged in more physical activity.

Out-of-sample Prediction. Moving beyond linear regression based analysis, we performed a further out-of-sample based prediction task by using two non-linear models: Random Forest and XGBoost. The aim was to develop models for predicting wearable health metrics, evaluating their performance during forcasting on unseen data. We used non-linear models to incorporate the more complex relationships between variables. Random forest is an ensemble learning method that operates by constructing a multitude of decision trees, whereas XGBoost is an ensemble of decision trees with gradient boosting that is able to ignore any vacuous features that may be present to prevent overfitting. For each non-linear model, we developed three regression models, one for each of our wearable health metrics, and validated them with a leave-one-out cross validation. For each model, we included the five social media metrics (i.e., $morning\_peak$, $night\_dip$, $morning\_larks$, $night\_owls$, and $positive\_affect$) and the two control variables (i.e., log_pop and log_gdp) as predictors. By inspecting the mean absolute error (MAE) values of all the six models (Figure 7), we observed results similar to those found through the previous linear regression models in Table 2. We can observe that XGBoost models generally outperformed the random forest models. Sleep quality was the easiest wearable health metric to predict, with an overall error of 8.76% for the best performing XGBoost model (Figure 7), followed by sleep quantity, with an error of 10.25%. Physical activity, on the other hand, yielded the highest mean absolute error of 19.24%, in line with what was observed with the previous linear regressions.

We also conducted an ablation study in which we predicted a linear combination of all those three health metrics but, instead of using the seven predictors, we used six of them by removing one predictor at a time. Removing any of the seven predictors yielded an error that went from 10.25% (when removing $night\_dip$) to a maximum error of 11.38% (when removing $morning\_larks$), confirming, again, the relative high importance of $morning\_larks$ as a predictive feature.

Summary of the Prediction Tasks. Overall, these results suggest that it was possible to predict sleep quantity and quality with the social media metrics, and that circadian factors were indeed associated with collective health. However, the results also indicate that physical activity was harder to predict than sleep, with our best performing model achieving an Adjusted R^2 of 0.492. This is in line with: *1)* circadian

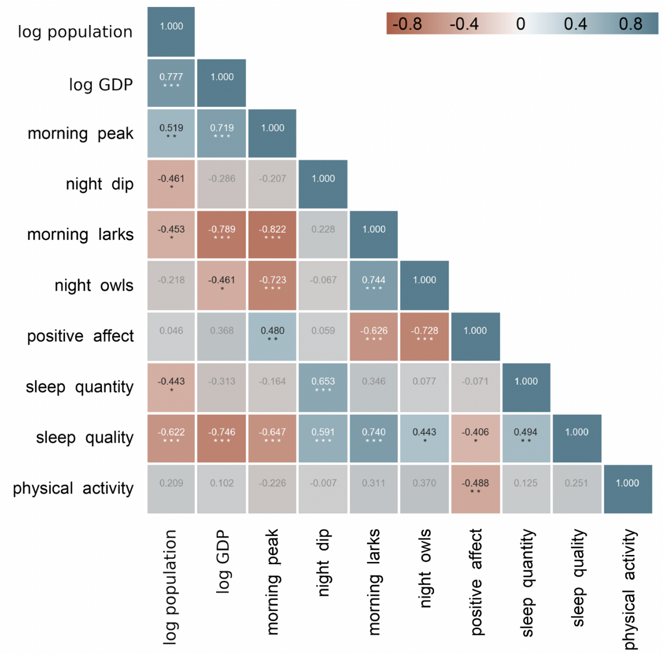


Figure 6: Cross-correlations among wearable health metrics and social media metrics. The correlations that are statistically significant are marked with a number of *'s based on their significance levels (i.e., $***p < 0.01$; $**p < 0.05$; $*p < 0.1$); those that are not ($p >= 0.1$) are colored in gray. Generally, health metrics positively correlated with each other: higher sleep quantity and quality also meant more physical activity. The prevalence of morning larks and night dips of online activity are the two social media metrics that correlated the most with sleep activity, whereas positive affect is the most correlated variable with physical activity.

factors being associated with sleep patterns more closely than general health, not least because, by definition, the circadian cycle is "the alteration of sleep and waking and all the bodily cycles attendant on those states" (Lynch 1972); and *2)* our previous observation that a city's social media posting is in-synch with sleeping patterns more than with generalized alertness. Furthermore, the prevalence of morning larks was a very strong predictor of sleep quantity, sleep quality, and physical activity, in line with prior studies detailing the health benefits that come with such a chronotype (Walsh et al. 1991; Reid, Maldonado, and Baker 2002; Biss and Hasher 2012).

6 Discussion and Conclusion

Main Results. The fact that quantity and quality of sleep are predicted by (over)-use of social networks in the early morning/late evening might be expected. Yet, beyond that, our results showed that "collective rhythms" on social media are associated with collective markers of sleep and physical activity, confirming small-scale studies (Walsh et al. 1991; Reid, Maldonado, and Baker 2002; Biss and Hasher 2012). More importantly, proxies for the otherwise inaccessible wearable readings could be extracted from publicly-available social media data, but they are so under previously unknown conditions that this study has now established: circadian rhythms can be reliably extracted late at night or during early morning, likely because, during the day, activity on social media tended to be more noisy (a variety of activities tend to be conflated at that time).

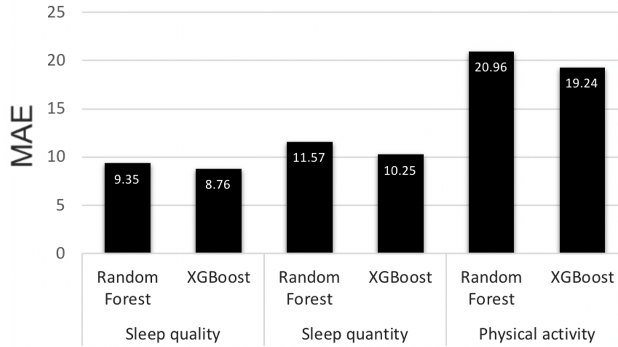


Figure 7: Mean absolute errors (MAE) for the out-of-sample prediction of sleep quality, sleep quantity, and physical activity with both random forest and XGBoost models. In a way similar to the results for the linear models, sleep quantity is the easiest to predict (with a 8.76% error for the best performing XGBoost model) and physical activity is the hardest (with a 19.24% error for the best performing XGBoost model).

Implications. Knowing people's circadian rhythms in a city is important for a number of reasons. It can inform decisions around lighting and noise levels in public spaces, as well as the scheduling of events and transportation. For example, if the majority of a city's population tends to be most alert and active in the morning, it may be more beneficial to schedule construction or other loud activities for the afternoon or evening. For public safety and security, knowing the natural patterns of sleep and wakefulness for a city's population can inform municipalities' decisions. For example, if most people in a city tend to be awake and active during the night, it may be necessary to increase patrols or lighting in certain areas to ensure public safety. Additionally, understanding circadian rhythms can inform the design of buildings and outdoor spaces, with the goal of creating environments that support the natural rhythms of the people who use them.

Limitations. Our work has two main limitations that call for future research. First, our findings hold for our 18 cities in developed countries, and may not be generalized to cities in more socio-economically deprived countries. Additionally, while our dataset spans across three continents, we do acknowledge that US and Asia are less represented compared to Europe. As commercial wearables will be widely adopted, analyses similar to ours could be repeated for a larger number of cities, or even for the entire globe, both boosting the statistical significance of our existing results and allowing for cross-cultural comparisons. Second, a variety of data biases should be considered when interpreting our results. Our cities represent westerns, mostly white cities, and are not stratified for basic covariates. Social media data may oversample techies (specifically, white male in their 20-30s). More generally, neither sets of our users represented a stratified sample of a city's inhabitants. On the one hand, our wearable users represented high-end consumers who are likely to enjoy specific lifestyles (the 18 cities under study have enjoyed high penetration rates of wearable devices over the recent years (Statista 2021)). On the other hand, our social media (Reddit) users are likely to be tech-savvy as well, but may only partly overlap with the set of wearable users. However, having two non-fully-overlapping sets of users (disjoint wearable/social media cohorts) was, in our case, a desirable experimental setup. Given the strong associations registered between the two sets, we established that our metrics captured the underlying behavioral markers of each city rather than those of the two specific samples. Finally, in the future, with sets of data more fine-grained than ours (e.g., data at the level of individuals, data geo-referenced at the level of neighborhoods, data of fine-grained sleep signals and other health indicators), researchers could focus on important within-city and within-subject similarities and differences. In addition, we would like to expand our study to understand trends at the country and culture level.

A Validation of Correlations Between Two Disjoint Datasets

In this paper, we used both Reddit and wearable data for our experiments while both datasets might represent different population groups. Ideally, one should calculate the correlations based on the same set of user cohorts from the two datasets to eliminate any potential bias, and correlations found due to chance. However, it is difficult to collect large-scale data and perform experiments in a scalable way that could reach multiple cities, if we were to collect both social media and wearable data for the same set of users. In addition, due to the sensitive nature of user data (e.g., privacy), we used aggregated data for the study and did not have access to the unique user characteristics data. For example, demographics information (e.g., age and gender) is not available on Reddit. Therefore, it is impossible to select a subset of 'matched users' based on user characteristics from the two datasets.

However, despite these limitations from the data perspective, we do believe that the correlations we found are not due to chance. This is because that, despite of having two non-fully-overlapping sets of users (disjoint wearable/social media cohorts), the prediction results demonstrate strong associations between those two sets. Our analysis was based on two relatively independent signals (i.e., sleeping patterns and physical activity). Due to this, we hypothesize that the probability of achieving significant correlations on both signals due to chance is quite low.

To validate further that these correlations were not due to chance, we conducted two experiments:

# user reassignment	0%	25%	50%	75%	100%
morning peak	-0.65	-0.19	-0.09	0.03	0.02
night dip	0.59	0.26	0.09	0.08	-0.03
morning larks	0.74	0.31	0.12	0.02	0.02
night owls	0.44	0.18	0.15	0.04	-0.04
positive affect	-0.41	-0.27	-0.12	-0.10	-0.08

Table 3: The change of average correlations between sleep quality (quantified by wearable data) and different circadian rhythm metrics (quantified by social media data) after reassignment of city users at different percentages.

- First, to have a sanity check, we randomly shuffled the circadian rhythm metrics at city level, and then correlated them to the collective health metrics. We bootstrapped this experiment for $1,000$ time, and calculated the percentage of trial experiments that found statistically significant correlations (with a p-value < 0.01). We found that the correlations over 99% of the trials were not statistically significant. This is an expected result, indicating that there are no intrinsic city characteristics that result in significant correlations between social media circadian rhythm metrics and wearable health metrics.
- Second, we aimed to understand whether we would still find significant correlations if we were to change the composition of users within the cities for one given data source. The hypothesis is that, if the correlations are by chance, by reassigning users across cities on one dataset, we would still find some significant correlations. Therefore, for each city, we randomly reassigned a certain percentage of its social media Reddit users to another city. We performed this experiment 10 times for each setting, and reported the average correlations between the selected social media based circadian rhythm metric and the selected wearable based health metric. We only report the results between the circadian rhythm metrics and the sleep quality wearable metric as we found very similar results for all metric combinations.

The results are shown in Table 3. We observed that, compared to the original correlations, the average correlations significantly deteriorate after reassigning city users. After reassigning more than 25% of Reddit users, the correlations between social media circadian rhythm metrics and wearable sleep quality metric are quite low (almost random). Both of those further experiments demonstrate that our obtained correlations were not due to chance.

References

Abdullah, S.; Matthews, M.; Murnane, E. L.; Gay, G.; and Choudhury, T. 2014. Towards circadian computing: "early to bed and early to rise" makes some of us unhealthy and sleep deprived. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 673–684.

Abdullah, S.; Murnane, E. L.; Matthews, M.; Kay, M.; Kientz, J. A.; Gay, G.; and Choudhury, T. 2016. Cognitive rhythms: Unobtrusive and continuous sensing of alertness using a mobile phone. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 178–189.

Aiello, L. M.; Quercia, D.; and Roitmann, E. 2018. Hearts and politics: Metrics for tracking biorhythm changes during brexit and trump. In *Proceedings of the ACM International Conference on Digital Health*, 111–115.

Althoff, T.; Hicks, J. L.; King, A. C.; Delp, S. L.; Leskovec, J.; et al. 2017. Large-scale physical activity data reveal worldwide activity inequality. *Nature*, 547(7663): 336–339.

Ancoli-Israel, S.; Cole, R.; Alessi, C.; Chambers, M.; Moorcroft, W.; and Pollak, C. P. 2003. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, 26(3): 342–392.

Balsamo, D.; Bajardi, P.; and Panisson, A. 2019. Firsthand opiates abuse on social media: monitoring geospatial patterns of interest through a digital cohort. In *The World Wide Web Conference*, 2572–2579.

Baumgartner, J. 2015. I have every publicly available Reddit comment for research. 1.7 billion comments 250 GB compressed. Any interest in this? https://redd.it/3bxlg7. Accessed: 2023-04-21.

Bettencourt, L. M. A.; Lobo, J.; Helbing, D.; Kühnert, C.; and West, G. B. 2007. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences*, 104(17): 7301–7306.

Biddle, J. E.; and Hamermesh, D. S. 1990. Sleep and the Allocation of Time. *Journal of Political Economy*, 98(5, Part 1): 922–943.

Biss, R. K.; and Hasher, L. 2012. Happy as a lark: Morning-type younger and older adults are higher in positive affect. *Emotion*, 12(3): 437.

Carrier, J.; and Monk, T. H. 2000. Circadian rhythms of performance: new trends. *Chronobiology International*, 17(6): 719–732.

Chaput, J.-P.; McNeil, J.; Despres, J.-P.; Bouchard, C.; and Tremblay, A. 2013. Seven to eight hours of sleep a night is associated with a lower prevalence of the metabolic syndrome and reduced overall cardiometabolic risk in adults. *PloS one*, 8(9): e72832.

Cho, K. 2001. Chronic 'jet lag' produces temporal lobe atrophy and spatial cognitive deficits. *Nature Neuroscience*, 4(6): 567–568.

Dos Reis, V. L.; and Culotta, A. 2015. Using matched samples to estimate the effects of exercise on mental health via twitter. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Facer-Childs, E. R.; Boiling, S.; and Balanos, G. M. 2018. The effects of time of day and chronotype on cognitive and physical performance in healthy volunteers. *Sports Medicine-open*, 4(1): 1–12.

Facer-Childs, E. R.; Middleton, B.; Skene, D. J.; and Bagshaw, A. P. 2019. Resetting the late timing of 'night owls' has a positive impact on mental health and performance. *Sleep Medicine*, 60: 236–247.

Foster, R. G.; and Wulff, K. 2005. The rhythm of rest and excess. *Nature Reviews Neuroscience*, 6(5): 407–414.

Golder, S. A.; and Macy, M. W. 2011. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051): 1878–1881.

Grandin, L. D.; Alloy, L. B.; and Abramson, L. Y. 2006. The social zeitgeber theory, circadian rhythms, and mood disorders: review and evaluation. *Clinical Psychology Review*, 26(6): 679–694.

Guntuku, S. C.; Buffone, A.; Jaidka, K.; Eichstaedt, J. C.; and Ungar, L. H. 2019. Understanding and Measuring Psychological Stress Using Social Media. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13, 214–225.

Hao, T.; Xing, G.; and Zhou, G. 2013. iSleep: Unobtrusive Sleep Quality Monitoring Using Smartphones. In *Proceedings of the ACM Conference on Embedded Networked Sensor Systems*, 1–14.

Jamison-Powell, S.; Linehan, C.; Daley, L.; Garbett, A.; and Lawson, S. 2012. "I can't get no sleep" discussing# insomnia on twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 1501–1510.

Karami, A.; Dahl, A. A.; Turner-McGrievy, G.; Kharrazi, H.; and Shaw Jr, G. 2018. Characterizing diabetes, diet, exercise, and obesity comments on Twitter. *International Journal of Information Management*, 38(1): 1–6.

Kountouris, Y.; and Remoundou, K. 2014. About time: daylight saving time transition and individual well-being. *Economics Letters*, 122(1): 100–103.

Kreitzman, L.; and Foster, R. 2011. *The rhythms of life: The biological clocks that control the daily lives of every living thing*. Profile Books.

Krejcar, O.; Jirka, J.; and Janckulik, D. 2011. Use of Mobile Phones as Intelligent Sensors for Sound Input Analysis and Sleep State Detection. *Sensors*, 11(6): 6037–6055.

Lin, H.; Jia, J.; Nie, L.; Shen, G.; and Chua, T.-S. 2016. What Does Social Media Say about Your Stress?. In *IJCAI*, 3775–3781.

Luyster, F. S.; Strollo, P. J.; Zee, P. C.; and Walsh, J. K. 2012. Sleep: a health imperative. *Sleep*, 35(6): 727–734.

Lynch, K. 1972. *What Time is this Place?* MIT Press.

McIver, D. J.; Hawkins, J. B.; Chunara, R.; Chatterjee, A. K.; Bhandari, A.; Fitzgerald, T. P.; Jain, S. H.; and Brownstein, J. S. 2015. Characterizing sleep issues using Twitter. *Journal of Medical Internet Research*, 17(6): e140.

Medvedev, A. N.; Lambiotte, R.; and Delvenne, J.-C. 2017. The anatomy of Reddit: An overview of academic research. In *Dynamics on and of Complex Networks*, 183–204. Springer.

Miller, M. A.; Rothenberger, S. D.; Hasler, B. P.; Donofry, S. D.; Wong, P. M.; Manuck, S. B.; Kamarck, T. W.; and Roecklein, K. A. 2015. Chronotype Predicts Positive Affect Rhythms Measured by Ecological Momentary Assessment. *Chronobiology International*, 32(3): 376–384.

Min, J.-K.; Doryab, A.; Wiese, J.; Amini, S.; Zimmerman, J.; and Hong, J. I. 2014. Toss'n'turn: Smartphone as Sleep and Sleep Quality Detector. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 477–486.

Moturu, S. T.; Khayal, I.; Aharony, N.; Pan, W.; and Pentland, A. 2011. Using social sensing to understand the links between sleep, mood, and sociability. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, 208–214. IEEE.

Murnane, E. L.; Abdullah, S.; Matthews, M.; Choudhury, T.; and Gay, G. 2015. Social (media) jet lag: How usage of social technology can modulate and reflect circadian rhythms. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 843–854.

Norlander, T.; Johansson, Å.; and Bood, S. Å. 2005. The affective personality: Its relation to quality of sleep, well-being and stress. *Social Behavior and Personality*, 33(7): 709–722.

Park, M.; Thom, J.; Mennicken, S.; Cramer, H.; and Macy, M. 2019. Global music streaming data reveal diurnal and seasonal patterns of affective preference. *Nature Human Behaviour*, 3(3): 230–236.

Park, S.; Constantinides, M.; Aiello, L. M.; Quercia, D.; and Van Gent, P. 2020. Wellbeat: A Framework for Tracking Daily Well-being Using Smartwatches. *IEEE Internet Computing*, 24(5): 10–17.

Reid, A.; Maldonado, C. C.; and Baker, F. C. 2002. Sleep behavior of South African adolescents. *Sleep*, 25(4): 417–421.

Richter, C. P. 1967. Sleep and activity: their relation to the 24-hour clock. *Research publications-Association for Research in Nervous and Mental Disease*, 45: 8–29.

Roenneberg, T.; Wirz-Justice, A.; and Merrow, M. 2003. Life Between Clocks: Daily Temporal Patterns of Human Chronotypes. *Journal of Biological Rhythms*, 18(1): 80–90.

Saha, K.; Bayraktaroglu, A. E.; Campbell, A. T.; Chawla, N. V.; De Choudhury, M.; D'Mello, S. K.; Dey, A. K.; Gao, G.; Gregg, J. M.; Jagannath, K.; et al. 2019. Social Media As A Passive Sensor In Longitudinal Studies of Human Behavior And Wellbeing. In *Extended Abstracts of the ACM CHI Conference on Human Factors in Computing Systems (CHI)*, 1–8.

Saha, K.; and De Choudhury, M. 2017. Modeling stress with social media around incidents of gun violence on college campuses. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW): 1–27.

Shameli, A.; Althoff, T.; Saberi, A.; and Leskovec, J. 2017. How gamification affects physical activity: Large-scale analysis of walking challenges in a mobile application. In *Proceedings of the ACM International Conference on World Wide Web (WWW) Companion*, 455–463.

Statista. 2021. Number of connected wearable devices worldwide by region from 2015 to 2022. https://www.statista.com/statistics/490231/wearable-devices-worldwide-by-region/. Accessed: 2023-04-21.

Tähkämö, L.; Partonen, T.; and Pesonen, A.-K. 2019. Systematic review of light exposure impact on human circadian rhythm. *Chronobiology International*, 36(2): 151–170.

Vink, J. M.; Vink, J. M.; Groot, A. S.; Kerkhof, G. A.; and Boomsma, D. I. 2001. Genetic Analysis of Morningness and Eveningness. *Chronobiology International*, 18(5): 809–822.

Walch, O. J.; Cochran, A.; and Forger, D. B. 2016. A Global Quantification of "Normal" Sleep Schedules Using Smartphone Data. *Science Advances*, 2(5): e1501705.

Walsh, J. K.; Schweitzer, P. K.; Anch, A. M.; Muehlbach, M. J.; Jenkins, N. A.; and Dickins, Q. S. 1991. Sleepiness/alertness on a simulated night shift following sleep at home with triazolam. *Sleep*, 14(2): 140–146.

Walsh, N. A.; Repa, L. M.; and Garland, S. N. 2021. Mindful larks and lonely owls: The relationship between chronotype, mental health, sleep quality, and social support in young adults. *Journal of Sleep Research*, e13442.

West, G. 2017. *Scale: The Universal Laws of Life and Death in Organisms, Cities and Companies*. Weidenfeld & Nicolson. ISBN 9780297869658.

Wittmann, M.; Dinich, J.; Merrow, M.; and Roenneberg, T. 2006. Social Jetlag: Misalignment of Biological and Social Time. *Chronobiology International*, 23(1-2): 497–509.

Zhang, Z. 2016. Variable selection with stepwise and best subset approaches. *Annals of Translational Medicine*, 4(7).