Hearts and Politics: Metrics for Tracking Biorhythm Changes during Brexit and Trump

Luca Maria Aiello  
Nokia Bell Labs  
luca.aiello@nokia-bell-labs.com

Daniele Quercia  
Nokia Bell Labs  
quercia@cantab.net

Eva Roitmann  
Nokia  
eva.roitmann@nokia.com

ABSTRACT
Our internal experience of time reflects what is going in the world around us. Our body’s natural rhythms get disrupted for a variety of external factors, including exposure to collective events. We collect readings of steps, sleep, and heart rates from 11K users of health tracking devices in London and San Francisco. We introduce measures to quantify changes in not only volume but also synchronicity and periodicity, and we empirically assess how strong those variations are, compared to random expectation, during four major events: Christmas, New Year’s Eve, Brexit, and the US presidential election of 2016 (Donald Trump’s election). While Christmas and New Year’s eve are associated with short-term effects, Brexit and Trump’s election are associated with long-term disruptions. Our results promise to inform the design of new ways of monitoring population health at scale.

1 INTRODUCTION
Our body pulses in cycles: we sleep or waken, are hungry or full, are alert or tired. The most dominant period in a person’s rhythms is the circadian cycle. Major departures from the normal range of the period have been associated with endogenous factors (e.g., illness) or exogenous ones (e.g., an external event inducing fear). Previous work has explored the relationship between circadian cycles and external factors, linking prolonged disruption of rhythms to pathological conditions, including cancer [18, 21]. Nowadays, “the alternation of sleep and waking and all the bodily cycles attendant on those states” [13] can be measured based on the use of social media [1, 6, 25], of augmented-reality games [3, 9], and, more reliably, of activity trackers [2, 19].

Yet, previous research has rarely ventured into: i) studying activity metrics beyond their volume; and ii) linking these metrics’ changes to global events. This study aims at explore these two aspects for the first time, and it does so by relying on large-scale data collected from Nokia Health monitoring devices used by 11,600 customers who live in London (67% of users) and San Francisco (33%) over the course of 1 year (from 1st April 2016 to 30th April 2017). Our users are 44% female, and their median age is 42 years. All users opted-in for research studies, and their data has been processed in an anonymized form. We consider three types of activities: total number of steps walked during the day; sleep duration measured in number of minutes slept at night; the estimated sleep time when the user went to bed for the night (hour and minutes, adjusted for timezone); and the average heart rate (beats per minute) measured once per day. Steps and sleep are measured by Nokia Health devices (e.g., former Withings wristbands and smart watches). All activities are measured at daily level for each user. For heart rates, when multiple measurements are available on the same day, we average them out. Our users represent a sample of the larger user population and are selected based on the number of days they used their devices: indeed, to reduce sparsity, we consider users who, for at least 90% of the days, measured their heart rates. This leaves us with 3.8M+ daily activity summaries.

By drawing from previous physiological and psychological studies, we derive metrics that relate walking, sleeping, and heartbeat to well-being. We characterize those three activities in terms of their volume (the raw value of the signal over time, §2), synchronicity (the degree to which the cycles of different people are in phase, §3) and rhythm (the activity periodicity, §4). We show how these metrics vary over the entire year, and how such variations represent distinctive signatures for four collective events: Christmas, New Year’s Eve, Brexit, and the US election of 2016. We find that users slept more than usual during the Christmas period and, as one expects, slept less than usual during Brexit, the US election, and New Year’s eve. Brexit and the US election are also associated with long-term disruptions in two main ways. First, in terms of sleeping patterns: users became heavily out-of-sync in the weeks after Brexit and even more so after the US election. Second, in terms of heart rates, we found major shifts in rhythm and volume, especially in the months around the US presidential election. These results suggest that our three metrics effectively capture how our biorhythms change during large-scale events, opening up new ways of monitoring population health at scale.

2 VOLUME
The amount of steps, hours of sleep and the dynamics behind heart rates have all been related to health outcomes. Physical activity boosts the levels of immune cells, and that results in a considerable reduction of sick days – from children to elderlies [15]. Sleep
deprivation has been found to make people accident-prone on the road, unproductive at work [27], and subject to brain aging [8, 27]. Sleep deprivation also increases the chances of ailments such as hypertension [23], cancer [5], diabetes, and obesity [16], and, as such, increases mortality rates [7]. In this work, to capture the amount of steps, sleep, and heart rates, given the measurement of an activity \( A \) (e.g., steps), we denote the activity of user \( u \) on day \( t \) with \( A_u(t) \), and compute the average activity during day \( t \) at population level as \( \bar{A}(t) = \frac{1}{|U|} \sum_{u \in U} A_u(t) \).

In Figure 1, we plot the average daily number of steps, hours of sleep, and heart rates for the whole year. The plots of steps and hours of sleep are spiky, and that comes from our typical weekly patterns: during weekends, people usually walk less and sleep more. Some of the spikes are much more prominent than the others though, and correspond to four major events (marked with dashed lines in the plots): the “Brexit” referendum in which the UK electorate voted to leave the EU on the 23rd of June 2016, the US presidential election on the 8th of November 2016, and Christmas and New Year’s Eve of the same year. For steps (first plot in Figure 1), there are two low points, which correspond to Christmas and New Year’s Eve. For hours of sleep (second plot), there are three low points, which correspond to Brexit, the US election, and New Year’s Eve. Finally, for heart rates (third plot), there are a few peaks and low points but they are limited – the most remarkable trend is represented by a considerable collective increase of heart rate just around the US election.

These results might suggest that increases in heart rates are in a cause-and-effect relationship with the US election. However, before considering causation, we need to rule out alternative explanations:

**New users.** If new users are suddenly introduced in the sample, heart rate volume could increase. That does not apply to our case since, for the whole duration of the year, we study the very same set of users whose heart rate is monitored almost continuously throughout the year (90%+ of the days).

**Software/hardware update.** Device and software updates might impact measurements. During the year of observation, our devices’ software that measured heart rates did not change, and all measurements showed high consistency.

**Physical Activity.** Heart rates could increase as a result of increased physical activity. However, there was no substantial change in daily number of steps at the time of the US election (a person did, on average, 6794 steps a day in October, 6750 in November, and 6660 in December).

**Temperature.** In cold weather, to keep the body warm, the heart beats faster. The temperature in the months concerning the US election was stable (Figure 2D), ruling out temperature as co-founding factor.

**Seasonality.** People’s rhythms are seasonal [1]. However the observed heart rate increases are steady and are not seasonal. If they were, given the comparable weather conditions (Figure 1), the heart rate levels in April 2016 would be the comparable to those in April 2017 – but they are not.

Upon observational data, it is hard to argue what caused heart rate increases. However, the strongest association appears to be with the US election, and that is because of three main reasons:

(i) **Alternative Explanations.** We have just ruled out the most plausible explanations other than the US election.

(ii) **External Validity.** Increases in heart rates have been found to be associated with emotional regulation and stress [10, 22, 24]. It should come as no surprise that the US election caused (self-reported) stress in a considerable part of the electorate. Based on a representative sample on 1,000+ US residents, a survey commissioned by the American Psychological Association found that more than half of the interviewees experienced the political climate around the presidential campaign as a significant source of stress [4].

(iii) **Dose-response relationships.** Dose-response patterns on observational data are necessary (but not sufficient) for considering causation. In our case, we indeed observe that events are strongly linked to biorhythm responses. To see how, contrast Londoners with San Franciscans: San Franciscans experienced rapid heart rate increases the last two months of the US political campaign (Figure 2C), experienced a peak exactly on the election day, and slept the least during the US election night (Figure 2B); by contrast, Londoners slept less the night after Brexit (Figure 2A), and started to experience heart rate increases on the US election day (Figure 2C), suggesting that their response was shifted compared to the US counterpart, as one expects. Therefore, dose-response relationships are observed for both the US election and Brexit.

### 3 SYNCHRONICITY

So far we have captured the volume of steps, hours slept, and heart rates. To go beyond volume, we now focus on temporal patterns.
The timing of behavior has always been a strong expression of the style of individuals and entire populations [13]. Nowadays that timing can be reliably captured by smart devices. Our sleep data, for example, includes the time at which users go to bed every day. This can be interpreted as an ordered sequence of timestamps, which is also called spike train. For the purpose of this study, we are interested in measuring the degree of synchronization between two users, that is, between two spike trains $s_1 = \{t_1^{(1)}, \ldots, t_n^{(1)}\}$ and $s_2 = \{t_1^{(2)}, \ldots, t_m^{(2)}\}$, within an interval $[0, T]$ of, say, one year. The SPIKE-distance function $D_S$ provides a parameter-free way of doing that [14]. It is defined as the integration of an instantaneous spike function $S(t)$ over time: $D_S(0, T) = \int_0^T S(t)dt \in [0, 1]$.

The spike function at time $t$ is defined as:

$$S(t) = \frac{|\Delta t_P(t)| \cdot (x_F(t))_n + |\Delta t_F(t)| \cdot (x_P(t))_n}{(ISI)_n}$$ (1)

where $\Delta t_P(t)$ is the difference between the two spikes $t_1^{(1)}(t)$ and $t_1^{(2)}(t)$ that immediately precede time $t$ in the two trains; $\Delta t_F(t)$ is the difference between the spikes following $t$; $x_F^{(n)}(t)$ is the distance between $t$ and the previous spike in the $n$th train; $ISI^{(n)}$ is the mean inter-spike interval in the $n$th train; and $(\bullet)_n$ denotes the average over the two trains. When $D_S = 0$, the two trains have no distance between them, meaning that their spikes are perfectly synchronized; when $D_S = 1$, the two trains are completely out-of-phase. The formulation of $D_S$ for the bivariate distance (for 2 users) can be extended to a multivariate case (for $2^+\text{ users}$) by averaging the distances of all the pairs of spike trains in the set. We compute that quantity and denote it with $\bar{D}_S$.

Each user’s sleep patterns for the entire year have been converted into a spike train. This consists of the sequences of times at which the user went to sleep. The level of de-synchronization in the population is then computed as the average spike distance score $D_S$ over all user pairs. Even if theoretically bounded in $[0, 1]$, the $D_S$ variable takes values from $0$ to $10^{-4}$ in our data (Figure 3A). That is because $D_S$ is quite low for events influenced by exogenous events (e.g., is it rare to find a considerable number of people who sleep in the middle of the afternoon). To quantify the extent to which synchronization changes after each of our four events, we compute the average spike distance $D_S$ over all user pairs in the week before the event, and in the week after the event. More formally, we define the Out-Of-Sync score ($OOS$) of an event $e$ occurring at time $t$ as $OOS_{e=\alpha} = D_S(t, t+\alpha) - D_S(t - \alpha, t)$, where $t$ is the time of the event $e$, and $\alpha$ is a buffer time window, which, in our case, we set to be of one week. If the event’s score is above zero, then this means that, after the event, the population became, on average, less synchronized. If it is below zero, then the population became more synchronized. To make sure that an event’s score is not due to chance, we contrast it to a null/random model, that is, we contrast it to what the score would be if computed at random days. More specifically, we compute $OOS$ with $\alpha$ of one week at 100 random days along the whole year, obtaining 100 scores. We then average those scores out to obtain the random model’s score, which is supposed to be zero. By definition, the accuracy of the $D_S$ measure (and, consequently, that of $OOS$) suffers in the presence of missing data points, which is the case for our data, since our devices are not perfectly reliable. As such, to get robust measurements, we filter out all users whose spike trains are not complete in the weeks before each event, and in the weeks after it.
This step turns out to exclude at most a few hundred individuals for each event.

After this filtering, we compute the out-of-synch scores \( \text{OOS} \). We find that, at random days, the scores are close to 0, as one expects. By contrast, the scores are subject to changes during three of our four events. More specifically, they do not change in a statistically significant way during Christmas, but they do considerably change during New Year’s Eve, Brexit, and the US election, suggesting that several users became out-of-synch after these three events (Figure 3B). To quantify the fraction of the population who slipped considerably out-of-synch after each event, we consider the frequency distribution of out-of-synch scores (Figure 3A): its right tail represents those user pairs who are heavily out-of-synch with each other. Using a standard practice in outlier identification [12], we consider all the points that are at least 2 standard deviations (2\( \sigma \)) as outliers: outliers(\( D_S(t, t + \alpha) \)) = \( \int_{\mu - 2\sigma}^{\mu + 2\sigma} f_{D_S(t,t+\alpha)}(x)dx \),

where \( \alpha \) is the considered time window (i.e., one week), and \( f \) is the probability density function computed for the variable \( D_S \) in the time period \([t, t + \alpha]\). To then measure the impact of an event \( e \), we compute the value for the previous expression of outliers() after \( e \) minus its value before \( e \), and normalize the result:

\[
\text{OOS}^r_{e@t} = \frac{\text{outliers}(D_S(t, t + \alpha)) - \text{outliers}(D_S(t, t - \alpha))}{\text{outliers}(D_S(t, t - \alpha))}
\] (2)

The resulting value is the Out-Of-Synch population growth (OOS\(^r\)): it is the relative increase in the portion of user pairs that are heavily out-of-synch. From Figure 3C, one sees that the value increased by 10% after New Year’s Eve and Brexit, and by as much as 30% after the US election. The random baseline shows no increase.

4 RHYTHMS

As a final metric, we consider circadian rhythm. This is a roughly 24 hour cycle in the physiological processes of living beings, including humans. Although circadian rhythms are endogenous ("built-in"), they are adjusted to the local environment by external cues such as light and temperature. Disruptions in a person’s circadian rhythm for sleep and heart rates have been found to have negative health consequences [20] and lead to pathological conditions [18, 21]. To see how to track circadian rhythms on our data of sleep patterns and heart rates, consider that any activity signal over time can be interpreted as a time series, an ordered sequence of activity measurements. To extract the period of an activity time series, one can use the Discrete Fourier Transform. This decomposes the temporal signal into a number of discrete frequencies which, if recombined, compose the original signal. The Power Spectral Density (PSD) is the distribution of relative power of those frequencies; we extract it using the Welch method [26]. To make the results more interpretable, we transform the frequencies of the PSD into periods (period(PSD)), which denote the amplitudes of the wave originated by those frequencies, expressed in number of days (e.g., a period of 7 days denotes a weekly pattern). More formally, given a user \( u \)’s time series in a period \([t_1, t_2]\), we define its characteristic rhythm as the period with the maximum PSD value in \([t_1, t_2] \): \( \text{rhythm}_u(t_1, t_2) = \arg \max(\text{period}(PSD(t_1, t_2))) \).

Our goal is to go beyond individual users and quantify the rhythm shift associated with an event in the entire population. To this end, for any given event \( e \) that took place at time \( t \), we first compute the rhythm shift an individual user \( u \) experienced before and after the event, within a temporal window \( \alpha \): rhythm shift\(_u@t \) = \( \text{rhythm}_u(t, t + \alpha) - \text{rhythm}_u(t, t - \alpha) \).

We then aggregate the rhythm shift values across all users by computing their frequency distribution \( f_{\text{rhythm shift}_{u@t}} \). To make sure our shift values are not due to chance, we resort to a null/random model. We compute such model’s score value by computing “rhythm shift” scores for 100 random days: we first compute individual shift scores around those days (rhythm shift\(_{u, \text{rand}@t} \)’), and then compute the distribution over all users and days (\( f_{\text{rhythm shift}_{u, \text{rand}@t}} \)).

Finally, to estimate the entire population’s rhythm disruption associated with \( e \), we compare the observed distribution with the distribution for random days: rhythm disruption\(_{e@t} \) = \( D_{\text{KL}}(f_{\text{rhythm shift}_{u@t}}, f_{\text{rhythm shift}_{u, \text{rand}@t}}) \). This is the KL divergence between the two frequency distributions [11]. The higher it is, the higher the rhythm shift that is associated with the event compared to random expectation (zero for no shift).

Figure 4 shows the distribution of rhythm shift for sleep and heart rates around Christmas and the US election. Each distribution is shown together with the corresponding null/random model’s distribution, and the difference between the observed distribution and the random one (called ‘rhythm disruption’) is also reported and denoted with “KL”. After Christmas and New Year’s Eve, the shifts for both sleep and heart rates are limited. Instead, after the US election and Brexit, the shift for sleep is considerable, and that for heart rates is disruptive\(^2\).

5 CONCLUSION

Based on all the results, one might hypothesize that each of our metrics could offer a way of profiling large-scale events. In reality, no individual metric considered in separation would be sufficient. For example, from Table 1, one can see that volume alone is not a

\(^2\)Due to page restrictions, we show the results for Christmas and the US election here and invite the reader to visit http://goodcitylife.org/health for more.
reliable marker for distinguishing the four events under consideration: Brexit is indistinguishable from New Year’s Eve, for instance. By contrast, considering our metrics in combination is sufficient for distinguishing the four events. Indeed, by plotting the daily average number of steps in a “rhythm disruption by volume” plot (Figure 5A), the four events are separable (i.e., they form distinctive clusters), suggesting that rhythm disruption and volume are, in our case, reliable markers for event classification. The same applies to daily average number of hours slept (Figure 5B). This is further supported by a DBSCAN clustering of those points, which returns a silhouette value (clustering quality value [17]) of $-0.35$.

Taken together, our results are very promising, yet three main limitations hold. First, our users are not representative of the general population. Second, our metrics suffer from data sparsity and, to be generalizable, they need to be further researched. Finally, our results do not speak to causation. Still, despite those limitations and based on the dose-response nature of the relationships between events and biorhythm measurements, we can conclude that, with our metrics at hand, one is able to capture “how we experience time” in unobtrusive ways. Synchronization and rhythms seem to be present in all living beings. They generally serve to keep the inner organisms working and keep the body coordinated with the external world. A failure of synchronization puts the body out-of-sync and under stress. Nowadays smart health trackers are able to capture that experience, and are able to do so at an unprecedented scale.

REFERENCES

[1] Scott A Golder and Michael W Macy. 2011. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. (2011).
[2] Tim Althoff, Jennifer L Hicks, Abby C King, Scott L Delp, Jure Leskovec, et al. 2017. Large-scale physical activity data reveal widespread activity inequality. Nature (2017).
[3] Tim Althoff, Ryen W White, and Eric Horvitz. 2016. Influence of Pokemon Go on physical activity study and implications. JMIR (2016).
[4] Sophie Bethune and Elizabeth Lewan. 2016. Stress in America. Technical Report. American Psychological Association.
[5] Scott Davis, Dana Mirick, and R G Stevens. 2001. Night Shift Work, Light at Night, and Risk of Breast Cancer. (2001).
[6] Munmun De Choudhury, Murali Kamar, and Ingmar Weber. 2017. Computational Approaches Toward Integrating Quantified Self Sensing and Social Media. In CSCW. ACM.
[7] Daniel F Kripke, Lawrence Garfinckel, Deborah L Wingard, Melvile R Klauber, and Matthew Marler. 2002. Mortality Associated With Sleep Duration and Insomnia. (2002).
[8] Jane Ferrie, Martin Shipley, Tasnime Akbaraly, Michael Marmot, Mika Kivimäki, and Archanza Singh-Manoux. 2011. Change in Sleep Duration and Cognitive Function: Findings from the Whitehall II Study. (2011).
[9] Eduardo Graeells-Garrido, Leo Ferres, Diego Caro, and Loreto Bravo. 2017. The effect of Pokemon Go on the pulse of the city: a natural experiment. EPJ Data Science (2017).
[10] Nis Hjortskov, Dag Rissén, Anne Katrine Blangsted, Nils Fallentin, Ulf Lundberg, and Karen Søgaard. 2004. The effect of mental stress on heart rate variability and blood pressure during computer work. European Journal of Applied Physiology (2004).
[11] S. Kullback and R. A. Leibler. 1951. On Information and Sufficiency. Annual Mathematics and Statistics (1951).
[12] Christophe Leys, Christophe Ley, Olivier Klein, Philippe Bernard, and Laurent Licata. 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology (2013).
[13] K. Lynch. 1972. What Time is this Place? MIT Press.
[14] Mario Mulansky, Nalbojsa Boznic, Andrea Strleia, and Thomas Kreuz. [n. d.]. A guide to time-resolved and parameter-free measures of spike train synchrony. In Event-based Control, Communication, and Signal Processing (EBCCSP). IEEE.
[15] David Nieman. 2000. Exercise effects on systemic immunity. (2000).
[16] John Reilly, J Armstrong, and A.R. Dorosty. 2005. Early life risk factors for childhood obesity: Cohort study. (2005).
[17] Peter J Rousseeuw. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics (1987).
[18] Saurabh Sahar and Paolo Sassone-Corsi. 2011. Regulation of Metabolism: The Circadian Clock dictates the Time. (2011).
[19] Ali Shameli, Tim Althoff, Amin Saberi, and Jure Leskovec. 2017. How Gamification Affects Physical Activity: Large-scale Analysis of Walking Challenges in a Mobile Application. In WWW. ACM.
[20] Joseph Takahashi, Hee-Kyung Hong, Caroline Ko, and Erin L McDearmon. 2008. The genetics of mammalian circadian order and disorder: implications for physiology and disease. (2008).
[21] Julian F Thayer, Fredrik Ahs, Mats Fredrikson, John J Sollers III, and Tor D Wager. 2012. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neuroscience & Biobehavioral Reviews (2012).
[22] Celine Vetter, Elizabeth E. Devore, Lani R. Wegryn, Jennifer Massa, Frank Speizer, Ichiro Kawachi, Bernard Rosner, Meir J. Stampfer, and Eva Schernhammer. 2016. Association Between Rotating Night Shift Work and Risk of Coronary Heart Disease Among Women. (2016).
[23] S. Kullback and R. A. Leibler. 1951. On Information and Sufficiency. Journal of the American Statistical Association (1951).
[24] Jane Ferrie, Martin Shipley, Tasnime Akbaraly, Michael Marmot, Mika Kivimäki, and Archanza Singh-Manoux. 2011. Change in Sleep Duration and Cognitive Function: Findings from the Whitehall II Study. (2011).
[25] Yafei Wang, Ingmar Weber, and Prasenjit Mitra. 2016. Quantified Self Meets Big Data: Sharing of Weight Updates on Twitter. In Digital Health. ACM.
[26] Peter Welch. 1967. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Transactions on audio and electroacoustics (1967).
[27] T. Wiseman. 2014. Night School: The Hidden Science of Sleep and Dreams. Pan Macmillan.