RESEARCH-ARTICLE

Biobehavioral Rhythms in Everyday Life: Data and Models for Capturing Cyclic Behavior in Naturalistic Settings

CHONG ZHAO, Grinnell College, Grinnell, IA, United States

MARIA ANA CARDEI, University of Virginia, Charlottesville, VA, United States

MATTHEW CLARK, University of Virginia, Charlottesville, VA, United States

RUNZE YAN, Emory University, Atlanta, GA, United States

AFSANEH DORYAB, University of Virginia, Charlottesville, VA, United States

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 9, Issue 4 (December 2025)
https://doi.org/10.1145/3770698
EISSN: 2474-9567

Biobehavioral Rhythms in Everyday Life: Data and Models for Capturing Cyclic Behavior in Naturalistic Settings

CHONG ZHAO*, Grinnell College, USA
MARIA CARDEI*, University of Virginia, USA
MATTHEW CLARK*†, University of Virginia, USA
RUNZE YAN, Emory University, USA
AFSANEH DORYAB, University of Virginia, USA

The ability to continuously and passively monitor human behavior using data from mobile and wearable devices in everyday environments presents new opportunities for tracking behavioral patterns that require detailed and long-term multimodal data. This paper introduces a dataset and corresponding methods for capturing and analyzing cyclical patterns of varying lengths from biobehavioral data collected passively through smartphones and wearable devices. The dataset includes up to 16 months of continuous records from smartphones and Fitbit devices, along with daily surveys from 166 university students. In addition to evaluating existing methods for modeling cyclical behavior, we also develop and present a new approach that facilitate multidimensional modeling and comparison of biobehavioral cycles within a population and across different time periods. We evaluate our methods using collected data to identify differences in cyclical behavioral patterns among various groups of students over different periods of the study.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing design and evaluation methods**; *Smartphones*; *Mobile devices.*

Additional Key Words and Phrases: Passive Sensing, Datasets, Modeling Cyclic Behaviors, Similarity, Physiology

ACM Reference Format:
Chong Zhao, Maria Cardei, Matthew Clark, Runze Yan, and Afsaneh Doryab. 2025. Biobehavioral Rhythms in Everyday Life: Data and Models for Capturing Cyclic Behavior in Naturalistic Settings. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 4, Article 241 (December 2025), 36 pages. https://doi.org/10.1145/3770698

1 INTRODUCTION

Human rhythms consist of repeating patterns that are inherent to human physiology and behavior. First characterized in 1976, biobehavioral rhythms encompass physiological, behavioral, and neurological cycles that influence daily life [52]. These rhythms can vary substantially across individuals and groups and fluctuate over time in response to changing contexts. Yet, despite their foundational role in shaping human behavior, biobehavioral rhythms remain understudied. A key barrier is the need for large-scale, fine-grained biobehavioral data to identify and analyze these human rhythms effectively. Smartphone and wearable devices now enable continuous,

*Authors contributed equally to this research.
†Corresponding author.

Authors' Contact Information: Chong Zhao, Grinnell College, Grinnell, Iowa, USA, zhaochon@grinnell.edu; Maria Cardei, University of Virginia, Charlottesville, Virginia, USA, cbr8ru@virginia.edu; Matthew Clark, University of Virginia, Charlottesville, Virginia, USA, wdd6ts@virginia.edu; Runze Yan, Emory University, Atlanta, Georgia, USA, runze.yan@emory.edu; Afsaneh Doryab, University of Virginia, Charlottesville, Virginia, USA, ad4ks@virginia.edu.

ACM 2474-9567/2025/12-ART241
https://doi.org/10.1145/3770698

longitudinal collection of naturalistic behavioral data in real-world settings, supporting investigations of how rhythms differ across contexts and populations. However, most existing datasets lack the combination of study duration, sensor sampling resolution, sample size, and multimodal integration needed for robust investigations of cyclical patterns. As a result, methods for modeling and comparing these rhythms are limited.

Although techniques such as cosinor analysis can identify periodicity [63], it remains unclear how rhythmic features vary across demographic groups, evolve over time, or relate to broader biobehavioral contexts. This gap limits our ability to systematically characterize rhythmic regularity, divergence, and disruption. In this paper, we take steps toward enabling such investigations by introducing new data resources and analytical approaches to study the regularity and variability of human rhythms across multiple timescales.

We conducted a naturalistic longitudinal study, collecting continuous, high-frequency passive sensing data and daily self-reports from 166 college students over a period of up to 16 months, spanning from September 2021 to January 2023. Participants lived their lives without intervention, allowing for the observation of naturally occurring cyclical patterns in behavior. The dataset was specifically designed to enable the modeling and analysis of cyclic behaviors, making it uniquely suited for studying human rhythms. Its large sample size, extended duration, and rich multimodal sensor data, which include physiological and behavioral measures from Fitbits [1], the smartphone-based AWARE application [25], and daily self-reported well-being, demographic, and personality information, provide an unparalleled foundation for rhythm analysis. Data were recorded at a high frequency, up to once per minute, providing a fine-grained temporal resolution. In addition, participants completed daily well-being surveys, offering detailed, longitudinal insights into individual and population-level behavioral rhythms. We provide the raw data and precomputed behavioral passive sensing and rhythm-related features to facilitate accessibility and downstream analysis.

We also introduce three novel metrics to quantify multimodal cross-sectional variability in rhythmic patterns. The *Rhythm Variability Score (RVS)* captures the rhythmic divergence across varying dimensions, within a fixed evaluation context for each fixed parameter. The *Aggregate Variability Score (AVS)* then aggregates these RVS values across evaluation contexts for each fixed parameter. Finally, these AVS values are averaged across all parameters to yield the *Mean Aggregate Variability Score (MAVS)*. We apply these metrics to the dataset, highlighting clear variability in rhythms between graduate and undergraduate students, various energy level-based groups, and between three key time windows during the fall semester. These findings highlight the value of these metrics in capturing nuanced differences in biobehavioral rhythms and underscore their potential to enhance our understanding of how human cyclic behavior varies across populations and temporal contexts.

The contributions of our work can be summarized as follows:

(1) We introduce a novel dataset designed to enable the extraction and modeling of cyclic patterns in human behavioral rhythms. Its extended duration, high temporal resolution, and multimodal nature allow for the observation of daily, weekly, and monthly cycles. Furthermore, the dataset enables the examination of how the regularity and characteristics of these rhythmic patterns relate to various contexts and internal psychological factors.
(2) We propose a novel approach for cross-sectional comparison of the extracted rhythms. We define three metrics that enable the comparison of rhythms across groups and dimensions. These metrics aggregate across rhythmic parameters and contexts to easily detect and analyze similarities and differences in rhythm patterns.
(3) We apply our method to our dataset to reveal biobehavioral rhythm variances between participants over time. From our analysis, we identify differences in rhythms between graduate and undergraduate students, between varying reported energy levels covering cognitive, emotional, and physical energies, and over three key weeks during the fall semester.

This paper presents related work in Section 2, followed by an overview of the dataset in Section 3. Detailed descriptions of the methods employed for cyclic pattern extraction and analysis are in Section 4. Resulting analyses, findings, and implications are found in Section 5. We discuss limitations and future work in Section 6. Finally, we conclude in Section 7 with reflections and final remarks.

2 BACKGROUND AND RELATED WORK

In this section, we first provide an overview of existing well-being-focused passive sensing applications and datasets most comparable to ours, highlighting their limitations for rhythm analysis (Section 2.1). We then summarize related work in modeling cyclic behaviors (Section 2.2).

2.1 Passive Sensing of Human Behavior

Passive sensing is a powerful tool for collecting naturalistic behavioral data without disrupting individuals' daily lives. Numerous studies have leveraged passive sensing streams such as smartphone activity, wearable sensor data, and environmental signals to infer well-being-related insights [33]. In particular, passively sensed behaviors have been connected to a myriad of health and wellness concerns. Several studies focus on mental health, highlighting how different behaviors can indicate that an individual is at risk for, or potentially experiencing, depression [23, 40, 61] and anxiety [26, 44, 51]. These passively sensed behaviors can also be applied to broader aspects of wellness, such as modeling stress [11, 58]. Sensing has also been used to power interventions for healthier behaviors, such as increased physical activity [13, 17, 36, 41] and healthier diets [46]. Despite investigating a number of applications, few studies have investigated the prolonged biobehavioral cycles underlying momentary behaviors.

In recent years, many datasets that leverage passive sensing have been made available to researchers. However, none of these datasets are sufficient for a robust study into human rhythms. Many existing datasets collected participants' data only over short periods, such as days [32, 56] or weeks [6, 8, 9, 31, 45, 58], limiting their ability to capture long-term cyclical patterns. While some datasets span longer periods, they have other limitations that restrict the study of human rhythms. For example, the Sleep Data dataset [20] collected sleep data over a period of eight years; however, it only includes unimodal data from a single participant and has not undergone peer review. In contrast, popular long-term datasets such as GLOBEM [62] and the College Experience [42] gather multimodal data over several years. Nonetheless, both studies rely on weekly self-reports rather than daily reports, which limits their capacity to examine the effects of short-term rhythms. Additionally, GLOBEM was not recorded continuously over the full four years; it only recorded participants' behaviors for 10 weeks each year, restricting researchers' ability to extract longer, month-long rhythms from the data. Additionally, while GLOBEM contains physiological data related to sleep, other important physiology readings, such as heart rate are not included. Similarly, the College Experience dataset lacks physiological data, which further limits the types of rhythms that can be analyzed. This dataset also provides sensor data at an infrequent sampling rate. While conditional sensing, such as phone unlocks and calls, is recorded whenever the event occurs, other behaviors, such as steps and activity recognition, are aggregated over larger periods, such as hours or days. With a less fine-grained sampling rate, the College Experience dataset cannot be used to comprehensively analyze behavioral rhythms, as shorter cycles cannot be detected. Our dataset aims to address these gaps by providing frequently sampled, multimodal behavioral observations over several months. A comparative overview of our introduced dataset alongside existing related datasets is provided in Appendix A.

2.2 Modeling Cyclic Behavior

Understanding human behavior can be enhanced by modeling cyclical patterns, with methods for conducting these analyses still evolving. Early efforts employed statistical techniques, such as Cosinor analysis, to model

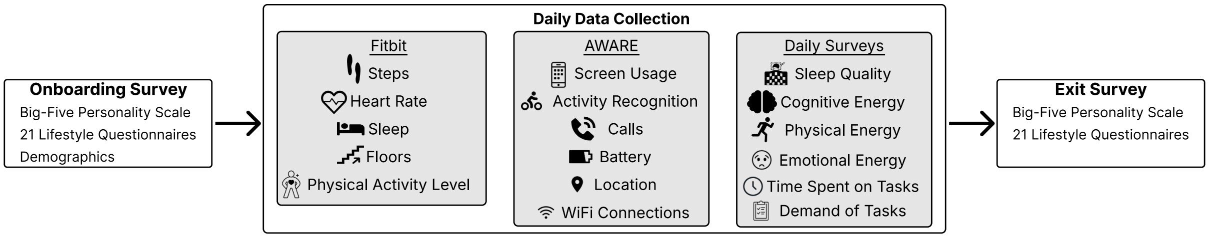


Fig. 1. The composition of the Human Rhythms Dataset. Participants completed a comprehensive survey consisting of 22 personality and lifestyle questionnaires during onboarding and after completing the study. Throughout the duration of the study, we collected data from provided Fitbits, the AWARE framework, and daily surveys.

biological and behavioral rhythms. This approach defines cycles using clear and interpretable rhythmic parameters [18]. Many traditional methods assume that cycles follow strict sinusoidal patterns, which can cause difficulties when trying to fit noisy or complex real-world signals. As a solution, more flexible approaches have emerged, such as CyHMMs, which utilize a cyclic hidden Markov model to reveal underlying cyclic structures without the need for a sinusoidal assumption [43]. Similarly, AutoNOM leverages changepoint detection to identify cycles, which it then uses to form sinusoidal functions [27].

Several studies have employed these methods to examine physiological rhythms and their relationship to participant wellbeing. Early research highlighted the impact of the circadian rhythm on sleep behavior [3], cognition, and alertness [2, 4]. The following studies leveraged Cosinor analysis to extract rhythmic parameters that describe the behavior and successfully linked these parameters to participants' depression and loneliness [63] and productivity [64]. These rhythmic patterns have also been utilized for more clinical applications, such as gauging a patient's readmission risk after pancreatic surgery [22]. Recently, large language models have been trained to parse through sensor data and summarize participants' daily routines [24, 67], offering promise for studying cyclical behaviors. Despite the value of these studies, many possible connections between human rhythms and participants' lifestyles remain uninvestigated.

3 THE HUMAN RHYTHMS DATASET

The Human Rhythms dataset contains passively sensed behaviors and daily survey responses from 166 college students at a mid-Atlantic American university collected over a 16-month period[1].

The median time that participants spent in our study was 224 days. Participants were compensated upwards of $15 per week for their data. This study was approved by the Institutional Review Board at our university. Details regarding data collection are provided in Figure 1 and in the following subsections. The released dataset includes a cleaned version of the raw data, as well as behavioral and rhythmic features extracted from it.

3.1 Data Collection

To protect participant privacy, all data sources anonymized the data as it was collected. During onboarding, each participant was provided with a three-character anonymous participant ID. Data from all sources was collected and immediately stored under this ID, enabling us to track data across the sensors and surveys without revealing the participant's identity. The true identity of each participant was maintained in a separate secure CSV file for compensation purposes. The research team began each onboarding session by explaining the nature of the study and requiring participants to provide informed consent by signing a digital consent form.

[1]Our dataset can be openly downloaded at https://github.com/HAI-lab-UVA/Human-Rhythms-Dataset.

Table 1. Selected demographics from our 166 participants.

Category	Count	Percentage	Category	Count	Percentage	Category	Count	Percentage	Category	μ	σ
Gender			Disability Status			Degree Pursuing			Age		
Female	109	65.66%	None	104	62.65%	Bachelor's	124	74.70%		20.57	3.68
Male	54	32.53%	Anxiety	48	28.92%	Doctor of Philosophy	30	18.07%	Years in the United States		
Non-Binary	3	1.81%	Depression	29	17.47%	Master's	9	5.42%		8.46	7.19
Race			ADHD	7	4.22%	Juris Doctorate	2	1.21%	Number of Roomates		
White	97	57.43%	Visual Impairment	5	3.01%	Medical Doctorate	1	0.60%		3.30	3.20
Asian	57	34.34%	ADD	3	1.81%	International Student					
Hispanic	56	33.73%	Other	6	3.61%	Yes	16	9.64%			
Black or African American	17	10.24%				No	150	90.36%			
Other	9	5.42%									

3.1.1 Onboarding and Exit Surveys. At the time of onboarding, participants filled out the *demographics* survey which included questions about gender, age, race, degree pursued, school at our University, marital status, language spoken at home, whether they were born in the United States or were citizens, living situation, employment status, and whether they were an international student, veteran, athlete, involved in Greek life, and a first-generation student. Finally, participants reported whether they had any diagnosed mental or physical disabilities. A summary of several demographics is provided in Table 1. Participants also filled out 22 baseline questionnaires at the beginning and the end of the study, providing information about personality and lifestyle. Each questionnaire is listed and briefly described in Appendix B.

3.1.2 Passive Sensing. Each participant received a Fitbit Sense [1] and installed the AWARE mobile application [25] to continuously record their behavioral and physiological data. The Fitbit tracked participants' steps, heart rate, activity level, and the number of floors climbed, with measurements taken every minute. We also collected sleep data, recording the times participants slept and their sleep stages, measured in seconds. Through the AWARE app, we collected information about participants' phone locations, WiFi connections, and activity recognition, sampling data from each sensor at intervals ranging from 30 seconds to 10 minutes[2]. Additionally, we collected data on participants' phone calls and screen time. These sensors are event-based, recording data only when specific actions occur, such as making/receiving a call or locking/unlocking the device.

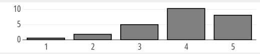
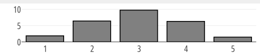
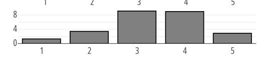
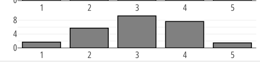
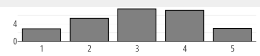
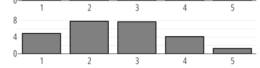
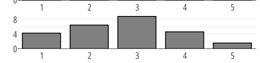
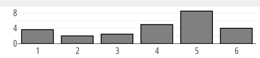
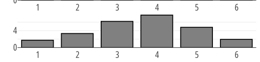
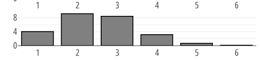
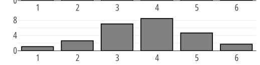
3.1.3 Daily Surveys. Each evening, participants were asked to complete a brief, 5-minute survey designed to gather ground truth measurements of their emotional, social, physical, and cognitive energy and activities. Participants used a 5-point Likert scale to rate their sleep quality and assess their cognitive, emotional, and physical energy levels throughout the day (Table 2). Each of these were measured with a single Likert scale to ensure the daily survey remained as short and unobtrusive as possible, ensuring participants remain willing to complete the survey each day. They also reported how much time they spent on various activities such as work, leisure, social interactions, and physical exercise. Finally, participants indicated how mentally and physically demanding their day was and how rushed they felt while completing their activities.

3.1.4 Additional Information. Our dataset focuses on passively sensed behaviors and self-reports from each participant. To help future research, our dataset also includes an "Additional Data" folder, which provides external events that may influence participants' behaviors. First, this folder contains a CSV file outlining all major events that occurred at our university, and may have altered behaviors. These events range from snow days and final exams to security events on campus. Similarly, we also provide a daily and hourly summary of the weather at our university. This data was collected through the Open-Meteo API[3] [68], and contains a variety of weather

[2]To protect participant privacy, our dataset does not include raw data for AWARE's WiFi, location, and application foreground sensors. Instead, we provide derived features from which participants cannot be identified.

[3]Weather data is publicly available under a CC-BY license. Data can be accessed at open-meteo.com and https://zenodo.org/records/14582479.

Table 2. The questions and responses to our daily survey. All response distributions are reported in the thousands.

Question	Response Options	Response Distribution
Sleep		
What was the overall level of your sleep quality?		
Energy Level		
What was the overall level of your physical energy?	1) Low 2) Somewhat low 3) Neither low nor high 4) Somewhat high 5) High	
What was the overall level of your emotional energy?		
What was the overall level of your cognitive energy?		
Daily Tasks		
How mentally demanding were your tasks?		
How physically demanding were your tasks?		
How hurried or rushed was the pace of your tasks?		
Activity Engagement Level		
How much did you engage in professional activities?	1) None 2) Less than an hour 3) 1–2 hours 4) 2–4 hours 5) 4–8 hours 6) More than 8 hours	
How much did you engage in social activities?		
How much did you engage in physical activities?		
How much did you engage in other leisure activities?		

data, including temperature, precipitation, wind, and cloud cover. Finally, since our study was run during the recovery phases of the COVID-19 pandemic, we also include the number of new COVID cases and deaths in our city, each day of the study. This was retrieved from the COVID-19 Repository by the Center for Systems Science and Engineering at Johns Hopkins University[4] [21].

3.2 Feature Extraction

We extracted features for downstream analysis. We first extracted traditional *behavioral sensor features* from each Fitbit and AWARE sensor using RAPIDS [57]. For each day that participants reported raw data, we extracted features to summarize the whole day, each hour of the day, each night (0:00 to 5:59), morning (6:00 to 11:59), afternoon (12:00 to 17:59), and evening (18:00 - 24:59). We also extracted features for several larger time scales, including each week, weekend, and sliding 3-day and 2-week window. A list of features and their descriptions can be seen in Table 3. We also extracted *rhythm parameters* [63] from our data using CosinorPy [39]. Using this function, we extracted the rhythm parameters displayed in Figure 2 and in Table 3. The methodology for extracting these features is explained in sections 4.1.1 and 4.1.2.

3.3 Post-Hoc Analysis of Data Completeness

After all data was collected and cleaned, we performed an initial validation of our data by analyzing the compliance from each data source. Participants were in the study for varying lengths of time, so we calculated the percentage

[4]COVID data publicly available under a CC-BY license. Data can be openly accessed at https://github.com/CSSEGISandData/COVID-19.

Table 3. A summary of features we extracted and included in the dataset. Behavioral sensor features were extracted using RAPIDS [57] for specific time periods. Due to limitations with iOS devices, some WiFi and applications features are only available for participants with Android devices ($n = 24$). Rhythm parameters were extracted from the entirety of participants' data using CosinorPy [39]. Our dataset also includes the raw data from most sensors, allowing future work to extract custom features. BI indicates a behavioral interpretation of the rhythm parameter.

Behavioral Sensor Features		Rhythm Parameters	
Heart Rate	Describe the range of heart rate values over the time period. Includes features such as minimum, maximum, entropy, average, and the number of minutes in various heart rate zones.	Period	Measures the duration of the oscillating function. BI: Duration of a full behavioral cycle (e.g., a 24-hour sleep-wake cycle).
Steps	Explore participants' steps throughout the provided time period, extracting features such as the sum of steps, timing of the first and last steps, and number of sedentary and active bouts.	Mesor	The midline of the oscillating function. If sampling rates are consistent, Mesor will equal the mean of data points on the function. BI: Baseline or average level of behavior across the cycle.
Sleep	Depict how well participants slept. Features include the amount of time participants spent asleep, and the time spent in each sleep stage.	Amplitude (Amp)	The difference between the mesor and the highest point along the oscillating function. BI: Strength of the rhythmic fluctuations (e.g., how much activity peaks differ from baseline).
WiFi	Describe the number of times the device scanned for connections or nearby devices, and the number of devices it found.	Magnitude	The difference between the highest and lowest points along the function. In single component Cosinor functions, the magnitude will always be double the amplitude. BI: Total range of behavioral variation within the cycle.
Screen	Monitor when participants lock and unlock their phone. These features explore the amount of time and number of interaction sessions participants had on their phone, both over the entire period and in each use session.	Acrophase (PHI)	The time at which the amplitude is reached, in relation to the start of the function. BI: Time of behavioral cycle when peak behavior occurs.
Location	Define participants' movement. These features include the amount of time participants spent at home, total distance traveled, average movement speed, and location entropy.	Orthophase	The difference in time between when the highest and lowest points along the function are reached. BI: Relative timing between the behavioral highs and lows.
Calls	Describe the number of phone calls participants made and received, and their length.	Bathyphase	The time at which the lowest point along the function is reached, in relation to the start of the function. BI: Time of day when behavior is at its lowest (e.g., least activity).
Battery	Explore the changes to the devices' battery level throughout the period. These features include the amount of time the battery spent charging and discharging, and the average battery consumption rate.	P-Value	The overall significance of how well the oscillatory function fits the provided data. BI: Indicates whether the detected rhythm is statistically significant.
Applications	Explain the time and number of sessions spent on different kinds of apps.	Residual Sum of Squares	A measure of the discrepancy between the observed data and the fitted model. A lower value indicates a better model fit. BI: Smaller values mean the rhythm model fits behavioral data better.
Activity Recognition	Describe the time participants spent stationary, walking, running, biking, and in a motorized vehicle.	Standard Error of Residuals (SER)	A measures the average distance between the observed data points and predictions by the fitted model. BI: Lower values indicate more stable and predictable behavior.
		Signal to Noise Ratio	The ratio between the rhythmic signal strength and background noise, indicating rhythm clarity and reliability. BI: Higher values reflect stronger and more distinct rhythms in behavior.
		Margin of Error	The uncertainty or confidence around the estimate rhythm parameters, reflecting model stability. BI: Lower margin of error = more reliable rhythm estimates.

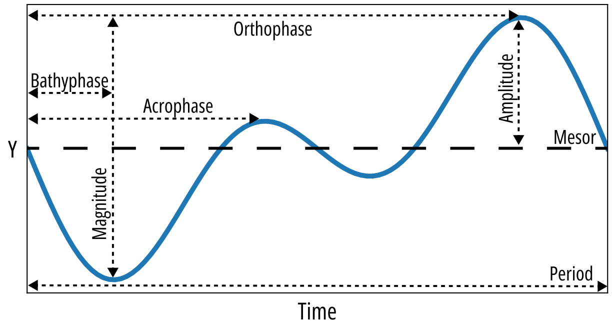


Fig. 2. Rhythm parameters, displayed on the cyclical function. Figure adapted from [64] with permission.

Table 4. A summary of the number of days of data collected from each data source. While exact amounts vary between data sources, we average over 200 days of sensor data and 150 days of survey data from participants.

	Mean	Standard Deviation	Minimum	Q1	Median	Q3	Maximum
Fitbit	223.31	94.42	11	200.25	224	237.75	491
Aware	200.6	97.83	3	146	212	232	478
Survey	151.79	68.91	4	114	160	192	419

of days with valid data from each sensor, and report the results in aggregate. Despite variation between data sources and participants, our participants were largely compliant with submitting data.

Most participants (56.63%) had between 200 and 250 days of sensor data. This distribution occurred because most participants were onboarded at the start of the school year and chose to leave the study at its conclusion, although this was not a requirement for anyone who was not graduating. Approximately a fifth (19.66%) of the participants opted to remain in the study and submitted more than 250 days of data. The distribution of days of data collected from each source is described in Table 4 and shown in Appendix C.

Fitbit Completeness. Missing Fitbit data may result from participants not wearing or syncing their devices. Depending on the sensor, Fitbit addresses missing data by either providing dummy data (placeholder values) or not supplying any data at all. We found valid data for approximately 87% of days for most sensors. However, some sensors had significantly less data. Valid sleep data was recorded for only 82.55% of days, suggesting that some participants likely did not wear their devices while sleeping. Similarly, the sensors measuring "fairly active" minutes (70.81%) and "very active" minutes (68.73%) showed substantial missing data, likely because participants needed to enter a specific heart rate zone for data to be recorded.

AWARE Completeness. Missing AWARE data indicates that participants either closed the app, ignored the notification to reopen it, or turned off their phones. The data completeness rates for location (92.68%) and screen status (93.70%) were higher than those for Fitbit. Meanwhile, WiFi and activity recognition both achieved a data completeness rate of 86.41%, which is roughly equal to that of the Fitbit data. Notably, the calls sensor in AWARE had the lowest completeness rate at 74.07%, as no data was recorded on days when participants did not make or receive any calls.

Table 5. Variables used to formalize our methods: single-period Cosinor model (left) and multidimensional rhythm variance (right).

Symbol	Description	Symbol	Description
$y(t)$	Single-period Cosinor model	RVS	Rhythm Variability Score
t	Observed time points	AVS	Aggregate Variability Score
P	Rhythm period	MAVS	Mean Aggregate Variability Score
β_1	Cosine coefficient	$R = \{r_1, \ldots, r_n\}$	Set of rhythm parameters, $\|R\| = n$
β_2	Sine coefficient	$S = \{s_1, \ldots, s_m\}$	Set of sensor features, $\|S\| = m$
A	Total amplitude	$F \in \{R, S\}$	Set of features the fixed feature f belongs to.
M	Mesor	$f \in F$	Fixed feature of interest, $f = r^*$ if $F = R$, else $f = s^*$
$\eta(t)$	Error term	F_{comp}	Complementary set of F, $F_{\text{comp}} = S$ if $F = R$, else $F_{\text{comp}} = R$
		$\|F_{\text{comp}}\|$	Cardinality of the complementary set
		$z_a \in F_{\text{comp}}$	a-th feature in the complementary set.
		D	Set of varying dimension combinations
		$d \in D$	Specific pairwise combination from D
		C_D	Total number of combinations of $d \in D$
		E	Set of evaluation contexts
		$e \in E$	Specific evaluation context from E
		$(\cdot)^*$	Targeted or fixed variable (generic form)
		p	P-value
		α	Significance

Survey Completeness. All 166 participants completed the onboarding demographic survey; however, two participants did not finish the pre-baseline survey. We found that 44 participants did not complete the post-baseline survey, resulting in a compliance rate of 73.49%. On average, participants completed 72.55% of the daily surveys they received.

4 METHODS

Our approach applies existing methods to detect significant rhythmicity and model these rhythms, then introduces a novel technique for analyzing multidimensional variance in cyclic behavior. We describe rhythm detection algorithms in Section 4.1, including periodicity detection (Section 4.1.1) and Cosinor-based modeling (Section 4.1.2), followed by our new approach for modeling and extracting insights from cyclic behaviors in Section 4.2. Table 5 covers key formulation notations.

4.1 Existing Rhythm Detection Algorithms

Previous work has established techniques for detecting significant periods and applying cyclical models, such as periodicity detection [63] and Cosinor cyclic behavior modeling [39].

4.1.1 Periodicity Detection. We use periodograms to detect the duration of significant periods, following the approach outlined in Yan et al. [63]. Specifically, we apply Fourier-based periodogram analysis [49] to each participant's feature data across all sensors. This method transforms time-domain signals into the frequency domain to detect statistically significant periods without requiring strong model assumptions. For each time series, we compute the spectral power across frequencies, apply a multiple-comparisons corrected threshold to identify significance, and extract peak periods corresponding to significant rhythms. This approach enables the detection of all periods that exhibit statistically significant rhythmicity without imposing strong assumptions, across different granularities such as minutes, hours, and days. Applied to the Human Rhythms dataset, it not only captures canonical rhythms such as the 24-hour, 12-hour, and 8-hour cycles reported in previous studies [3, 63], but also allows for the exploration of additional periodicities, such as those associated with biorhythmic patterns, that may be present in individuals' data.

4.1.2 Cyclic Behavior Modeling. For each significant period, we model sensor rhythms using the Cosinor model, a standard approach for known-period rhythms [18, 28]. Intuitively, the Cosinor model fits a sine–cosine wave to the data, much like modeling daily or weekly cycles, making it straightforward to capture regular patterns (Figure 2). The single-period Cosinor model [39] is expressed as the following:

$$y(t) = \beta_1 \sin\left(\frac{2\pi t}{P}\right) + \beta_2 \cos\left(\frac{2\pi t}{P}\right) + M + \eta(t).$$

Here, t represents the observed time points in the time series, P is the known period of the rhythm, β_1, β_2 are the sine and cosine coefficients, respectively, M is the mesor, and $\eta(t)$ is the error term. The total amplitude of the rhythm is based on the sine and cosine coefficients $A = \sqrt{\beta_1^2 + \beta_2^2}$. By fitting this sinusoidal curve to the data, we estimate *rhythm parameters* such as total amplitude, acrophase, mesor, and related measures (see Table 3 for definitions and behavioral implications).

4.2 Analyzing Multidimensional Variance in Cyclic Behavior

To enable multimodal cross-sectional and cross-group assessments of rhythm variability, we developed an approach consisting of the Rhythm Variability Score (RVS) and an aggregation that consists of the Aggregate Variability Score (AVS) and the Mean Aggregate Variability Score (MAVS). These metrics capture and aggregate variability across contextual dimensions, evaluation contexts, and rhythmic parameters to provide a comprehensive view of rhythm divergence. The Rhythm Variability Score (RVS) quantifies the degree of rhythmic divergence by measuring the proportion of rhythm parameters that exhibit statistically significant differences across varying dimensions, *within a fixed evaluation context for each fixed parameter.* The Aggregate Variability Score (AVS) then aggregates these RVS values *across evaluation contexts for each fixed parameter* to provide a broader assessment of rhythmic dissimilarity. Finally, these AVS scores are averaged *across all parameters* to yield the Mean AVS (MAVS). This approach enables analysis of the stability and variability of cyclic behavior by first quantifying how rhythmic patterns differ across varying dimensions (e.g., time windows or groups) within a fixed context, and then aggregating these results across all such contexts to assess global rhythmic divergence.

4.2.1 Rhythm Variability Score. Let $R = \{r_1, r_2, \ldots, r_n\}$ denote the set of n rhythm parameters, and let $S = \{s_1, s_2, \ldots, s_m\}$ denote the set of m sensor features. Our objective is to quantify rhythm variability across rhythm parameters (e.g., mesor) and sensor features (e.g., mean number of steps), and to evaluate how these rhythmic patterns vary across multiple dimensions (e.g., demographic groups). To achieve this, we compute the Rhythm Variability Score (RVS) for a fixed feature of interest f, where $f \in R$ (rhythm-centric view) or $f \in S$ (sensor-centric view), by comparing it against every element in the complementary set F_{comp}. The set F_{comp} represents sensor features if the feature of interest is a rhythm parameter, or the set of rhythm parameters if the feature of interest is a sensor feature. This comparison is done across a set of varying comparison contexts D (e.g., groups), where each $d \in D$ represents a specific pairwise combination. This is further performed within a fixed evaluation context $e^* \in E$ (e.g., one week). Both D and E are contextual dimensions derived from the data, drawn from the same underlying distribution of contextual factors (e.g., demographic groups or time windows), but they represent distinct axes of comparison and evaluation within a given computation. For example, if D denotes varying demographic groups across which rhythmic differences are assessed, then E may represent time windows, with the evaluation fixed within a specific time window e^*; conversely, if D varies across time windows, then E may represent demographic groups, with e^* denoting a fixed group context.

To identify statistically significant differences between contextual levels (e.g., demographic groups) d, we perform a one-way ANOVA for each combination of rhythm parameters and sensor features. This test will identify whether any sensor feature or rhythm parameters differ between the conditions. We use the standard indicator function $1\{p \leq \alpha\}$ (equal to 1 if $p \leq \alpha$, 0 otherwise) to count significant results. The RVS is then defined as follows:

$$\mathrm{RVS}_f^{(D,e^*)} = \frac{1}{|F_{\mathrm{comp}}|} \sum_{a=1}^{|F_{\mathrm{comp}}|} \left(\frac{1}{C_D} \sum_{d \in D} \mathbf{1}\{p_{f z_a d}^{(e^*)} \leq \alpha\} \left(1 - \sum_{d \in D} p_{f z_a d}^{(e^*)} \mathbf{1}\{p_{f z_a d}^{(e^*)} \leq \alpha\} \right) \right), \tag{1}$$

$$C_D = \sum_{2 \leq b \leq |D|} \binom{b}{|D|}$$

where

- $f \in \{s^*, r^*\}$ is the fixed feature of interest, representing either a target sensor feature or a rhythm parameter ($f \in F$),
- F_{comp} is the complementary set to F, representing all sensor features if f is a rhythm parameter or all rhythm parameters if f is a sensor feature,
- z_a is the a-th element in the complementary set ($z_a \in F_{\mathrm{comp}}$),
- D is the set of varying dimensions or contexts (e.g., group combinations),
- $d \in D$ is a specific pairwise comparison between contextual levels drawn from D,
- C_D is the total number of valid combinations of elements from D,
- $e^* \in E$ is a fixed evaluation context (e.g., one week), and
- $p_{f z_a d}^{(e^*)}$ is the p-value associated with the fixed feature f, complementary feature z_a, and varying dimension d, computed within the fixed evaluation context e^*.

Equation 1 defines a template form of the RVS, which captures rhythmic divergence for a fixed feature f across a set of varying dimensions D, within a fixed evaluation context $e^* \in E$. This formulation aggregates the proportion and strength of statistically significant differences based on p-values and indicator functions. A higher RVS value indicates greater rhythmic divergence across the varying dimension D, as the score increases with both the number of statistically significant differences (captured by the indicator function) and the magnitude of those differences (reflected by the corresponding p-values). Accordingly, RVS quantifies whether rhythmic patterns differ across conditions and the degree to which they diverge. We can further instantiate this general formulation for two specific cases: 1) the sensor-centric case, where $f = s^* \in S$ and we aggregate over rhythm parameters R, and 2) the rhythm-centric case, where $f = r^* \in R$ and we aggregate over sensor features S. These are derived in Appendix D.

4.2.2 Aggregating and Averaging Variability Scores. To evaluate overall rhythmic divergence across multiple evaluation contexts, we aggregate the RVS computed for each evaluation context $e \in E$. This aggregation yields the Aggregate Variability Score (AVS) for a fixed feature f. For example, if we are comparing a rhythm's mesor between weeks, then AVS would be calculated by aggregating each week's mesor RVS. Individual AVS_f scores can then be averaged across all fixed features $f \in F$, where F denotes either the set of rhythm parameters R or sensor features S, to compute the Mean Aggregate Variability Score (MAVS). The calculation for AVS and MAVS are derived below, with a directional arrow indicating the transition from AVS to MAVS:

$$\mathrm{AVS}_f = \frac{1}{|E|} \sum_{e \in E} \mathrm{RVS}_f^{(D,e)} \quad \Rightarrow \quad \mathrm{MAVS}_F = \frac{1}{|F|} \sum_{f \in F} \mathrm{AVS}_f \tag{2}$$

Equation 2 captures how consistently a specific rhythm parameter or sensor feature exhibits divergence across a set of comparison dimensions D under various evaluation conditions. The MAVS then averages these AVS values across the entire set of rhythm parameters R or sensor features S. This yields an overall rhythm variance score that reflects global rhythmic divergence at the population level, accounting for variability across all features and all evaluation contexts. Higher MAVS values indicate greater rhythm variability across groups or time, depending on the comparison. Since MAVS is derived from p-values, it reflects the extent of statistically significant differences. While its magnitude may vary with the number of features, MAVS is primarily used for relative comparisons across groups or time windows. We can further instantiate the general formulations of AVS and MAVS for both rhythm-centric and sensor-centric cases, detailed in Appendix D.

5 ANALYSIS

We summarize our analysis results in this section. Utilizing the Human Rhythms dataset, we first explore periodicity detection in Section 5.1, the impact of external events on rhythms in Section 5.2, followed by analysis of multimodal variance in cyclic behavior in Section 5.3. In Section 5.4 we dive deeper into analyzing differences between groups across time, and in Section 5.5 we investigate rhythm variability between time windows across group contexts. Finally, in Section 5.6, we validate our approach with an existing dataset.

5.1 Periodicity Detection

We extract the length of significant periods contained within the participants' behavior using Fourier-based periodogram analysis. We specifically recorded the percentage of participants whose data exhibited a significant cycle of that approximate length (rounded to the nearest hour) for selected behaviors. Figure 3 illustrates the number of participants who had statistically significant periods of various lengths. The high-resolution data enables the detection of short periods, of a single hour or shorter, to be extracted. Moreover, continuous long-term data collection leads to the presence of long-term rhythms that span several hundred days. Although these longer periods are present within the data, they tend to be specific to individual participants.

Analyzing periods that are identified within several participants' data provides insight into human behavior. As shown in Figure 3A, many behaviors frequently demonstrate 12- and 24-hour periods, as would be expected from daily behaviors. Interestingly, participants' step rhythms also appear to follow 8-hour periods, potentially aligning with participants' inactivity while sleeping each night. Other behaviors frequently follow hourly periods, which are more behavior-specific. In particular, battery consumption rate tends to follow periods of approximately 18 hours, potentially showing how participants charge their phones overnight. Similarly, participants' time sleeping tends to abide by 10-hour periods, slightly longer than the recommended 8 hours of sleep nightly, but still within the expected range for college students.

The daily periods (Figure 3B) also provide insight into the behavior of the participants. 24-hour behaviors are most common for all behaviors, capturing participants' daily routines. Longer periods for some behaviors, namely sleep and phone usage, are less common, emphasizing that these behaviors are mostly influenced by daily routine. 7-day periods are also relatively common among participants, demonstrating a clear weekly behavior, such as walking to class at the same time each day. Other behaviors, such as median heart rate, have many days with significant periods. Although 1, 2, 4 and 7 days are the most common, many participants have significant periods of different lengths, potentially indicating the presence of unstudied rhythms. Studying these periods for each participant may help enable context-aware interventions, personalized recommendations, and early detection of rhythm disruptions.

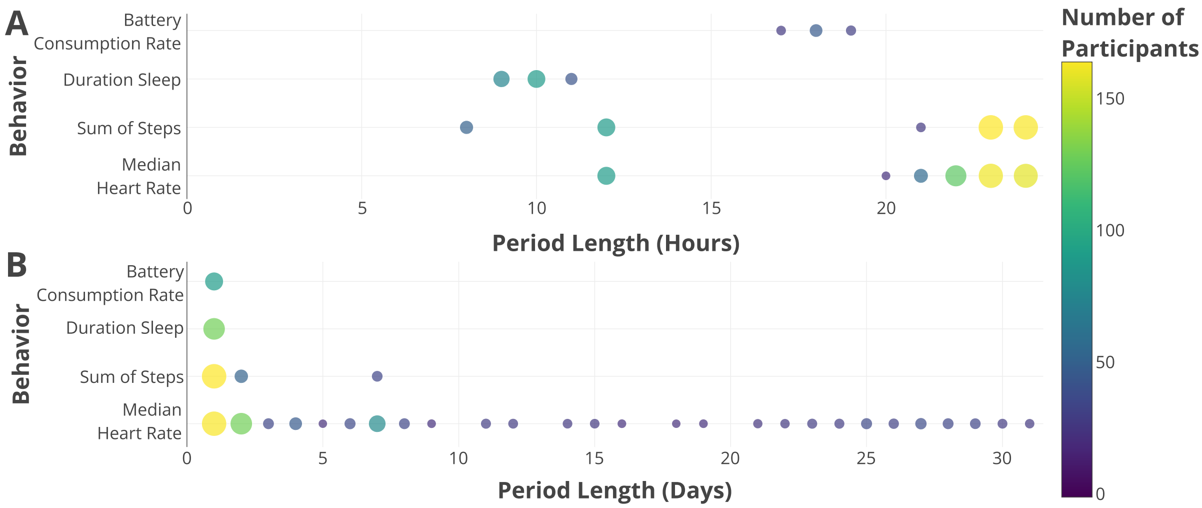


Fig. 3. The number of participants with a significant period, for all periods of a small subset of behavioral features demonstrated by at least 25% of participants, truncated to the corresponding hour (A) or day (B). The period lengths provide insight into participants' cyclical behavior.

5.2 Measuring the Impact of External Events

For further rhythm analysis, we investigate the impact of external events on rhythmic patterns. We segmented the data into three academic-calendar periods (fall semester, winter break, and spring semester) to capture distinct external contexts likely to influence participants' rhythms. Rhythm parameters from the Cosinor model were computed for each time range and sensor feature using a 24-hour period. Paired t-tests across participants identified significant changes in rhythm parameters between periods, revealing rhythm alterations and the proportion of features affected by external events.

Table 6 highlights significant rhythm shifts across academic periods, with the largest percent changes observed between the fall semester and winter break and again between winter and spring. Opposite signs of change across these transitions suggest that rhythms are likely relatively stable during semesters but disrupted during winter break. For example, mesor (baseline levels) increased for activity and location features during winter, while amplitude rose for location and battery features, reflecting reduced mobility and altered routines when students return home. These patterns reversed in spring, indicating re-stabilization of rhythms. Changes in the Standard Error of Residuals (SER) further show that battery usage, heart rate, WiFi, and step rhythms became more regular during winter, while location and mobility features grew more irregular. These findings demonstrate that external events can substantially alter daily rhythms and highlight behavioral signatures of disruption and recovery. For the ubiquitous computing community, such insights enable context-aware applications, such as adaptive health interventions, automated routine monitoring, and anomaly detection systems that account for external events (e.g., holidays or moves) when interpreting behavioral data.

5.3 Multimodal Variance in Cyclic Behavior

We examine the variance in cyclic behavior across multiple modalities and evaluation contexts using the variability metrics we define in Section 4.2. Specifically, we assess how rhythmic patterns vary across both group and temporal

Table 6. Rhythm variance in response to external events across different time periods. The parameter SER stands for standard error residual. The "Fall vs. Winter" and "Winter vs. Spring" columns show the average percentage change in each rhythm parameter across all participants for the specified sensor. Positive values indicate increases, while negative values indicate decreases. The percentages in parentheses represent the average absolute magnitude of these changes, regardless of direction. The % Feature shows the percentage of features that have a significant p-value for the paired t-test.

Parameter	Behavioral Features	Fall vs Winter	% Feature ($p < 0.05$)	Winter vs Spring	% Feature ($p < 0.05$)
Mesor	Activity Recognition	6.85% (13.04%)	83.3	-5.16% (12.38%)	83.3
	Battery	-10.97% (16.60%)	0.0	1.42% (11.36%)	16.7
	Heart Rate	-2.49% (2.49%)	100.0	2.32% (2.32%)	100.0
	Locations	50.47% (64.83%)	90.0	-9.95% (50.86%)	85.0
	Step	-12.50% (26.25%)	89.7	-34.94% (61.26%)	72.4
	Wifi Connections	-10.36% (10.36%)	50.0	6.42% (11.09%)	50.0
	Wifi Visible	-47.54% (47.54%)	50.0	43.96% (43.96%)	50.0
Amplitude	Activity Recognition	-16.70% (21.53%)	50.0	-10.65% (34.33%)	83.3
	Battery	114.90% (114.90%)	16.7	-72.03% (72.03%)	0.0
	Heart Rate	-17.14% (17.14%)	50.0	9.13% (18.75%)	40.0
	Locations	16.17% (36.90%)	20.0	-27.56% (53.22%)	80.0
	Step	-26.46% (38.48%)	86.2	59.37% (66.45%)	72.4
	Wifi Connections	-40.49% (40.49%)	50.0	82.11% (82.11%)	50.0
	Wifi Visible	-39.18% (39.18%)	0.0	65.41% (65.41%)	0.0
Acrophase	Activity Recognition	6.29% (11.28%)	33.3	-5.45% (13.20%)	33.3
	Battery	6.76% (14.82%)	0.0	14.83% (22.86%)	16.7
	Heart Rate	-2.54% (3.91%)	0.0	-3.76% (9.94%)	0.0
	Locations	1.17% (6.55%)	10.0	-0.01% (8.85%)	20.0
	Step	8.82% (15.85%)	55.2	-8.03% (16.90%)	65.5
	Wifi Connections	-1.44% (2.79%)	0.0	-3.28% (5.35%)	0.0
	Wifi Visible	-22.16% (22.16%)	0.0	-8.93% (8.93%)	50.0
SER	Activity Recognition	4.85% (9.16%)	66.7	-4.48% (10.53%)	83.3
	Battery	-16.20% (16.48%)	0.0	5.90% (5.90%)	33.3
	Heart Rate	-3.71% (6.26%)	90.0	7.44% (7.44%)	90.0
	Locations	31.78% (36.23%)	90.0	-15.66% (23.00%)	75.0
	Step	-9.83% (17.59%)	75.9	17.62% (29.68%)	82.8
	Wifi Connections	-21.75% (21.75%)	50.0	29.21% (29.21%)	100.0
	Wifi Visible	-27.14% (27.14%)	0.0	11.18% (11.18%)	0.0

dimensions, evaluating changes over rhythm parameters and sensor features. We first compare groups across time windows, and then reverse the comparison to assess differences between time windows within each group.

In this analysis, we consider two types of groupings: (1) demographic categories (e.g., undergraduate vs. graduate students) and (2) energy-based categories. The energy-based categories are derived by calculating weekly averages of participants' daily self-reported cognitive, emotional, and physical energy levels. Participants are categorized independently each week based on whether their average score is above or below the weekly population mean. We focus on three representative weeks from the fall semester—early (week 5), midterms (week 10), and finals (week 17)—to examine temporal variation in rhythms across groups. Participants with less than 20% missing data per target week were retained after interpolation-based imputation. For each group and week, single-subject 24-hour Cosinor models were fitted to support subsequent rhythm variability analyses. Table 7 reports results for RVS, AVS, and MAVS, and Table 8 does the same, but conversely aggregates across sensor features for all rhythm parameters. In both tables, the left-hand side tables report results for analyses between groups across time windows (BGAT), and the right-hand side shows variability metrics for analyses between time windows across groups (BTAG). Both paradigms are discussed in subsequent sections.

5.4 Rhythm Variability Between Groups Across Time (BGAT)

To compare rhythmic patterns between groups across time windows, we first compute the Rhythm Variability Score (RVS), which quantifies variability by measuring the proportion of rhythm parameters that show statistically

Table 7. Rhythm Variability Scores (RVS) are calculated between groups for each week (left-BGAT) and between weeks for each group (right-BTAG), for each target rhythm parameter r^* across all sensor features. The Aggregate Variability Scores (AVS) then aggregate RVS values across all weeks (BGAT) and groups (BTAG) for a given rhythm parameter. Finally, Mean AVS (MAVS) is calculated by averaging AVS values across all rhythm parameters. Higher RVS values indicate greater rhythmic divergence between groups in a given week (BGAT) and between weeks in a given group (BTAG), while higher AVS and MAVS values reflect greater overall variability across time (BGAT) and groups (BTAG) and rhythm parameters, respectively. Three key rhythm parameters are shown, which make up the basis of the Cosinor model.

Between Groups Across Time Windows (BGAT) for All Sensor Features				
Undergraduate vs. Graduate	Week 5	Week 10	Week 17	AVS_{r^*}
Mesor	0.33	0.41	0.04	0.26
Amplitude	0.24	0.25	0.08	0.19
Acrophase	0.16	0.38	0.21	0.25
$MAVS_R$				0.23
High vs. Low Cognitive Energy	Week 5	Week 10	Week 17	AVS_{r^*}
Mesor	0.10	0.17	0.0	0.09
Amplitude	0.22	0.14	0.04	0.13
Acrophase	0.10	0.18	0.04	0.11
$MAVS_R$				0.11
High vs. Low Emotional Energy	Week 5	Week 10	Week 17	AVS_{r^*}
Mesor	0.11	0.17	0.07	0.12
Amplitude	0.18	0.10	0.10	0.13
Acrophase	0.01	0.00	0.07	0.07
$MAVS_R$				0.11
High vs. Low Physical Energy	Week 5	Week 10	Week 17	AVS_{r^*}
Mesor	0.21	0.21	0.0	0.14
Amp	0.33	0.12	0.04	0.16
Acrophase	0.02	0.02	0.19	0.08
$MAVS_R$				0.13

Between Time Windows Across Groups (BTAG) for All Sensor Features			
Week 5 vs. Week 10 vs. Week 17	Undergraduate	Graduate	AVS_{r^*}
Mesor	0.53	0.001	0.27
Amplitude	0.43	0.02	0.23
Acrophase	0.28	0.01	0.15
$MAVS_R$			0.22
Week 5 vs. Week 10 vs. Week 17	High Cognitive Energy	Low Cognitive Energy	AVS_{r^*}
Mesor	0.47	0.41	0.44
Amplitude	0.39	0.34	0.37
Acrophase	0.09	0.23	0.16
$MAVS_R$			0.32
Week 5 vs. Week 10 vs. Week 17	High Emotional Energy	Low Emotional Energy	AVS_{r^*}
Mesor	0.53	0.41	0.47
Amplitude	0.33	0.36	0.34
Acrophase	0.22	0.08	0.15
$MAVS_R$			0.32
Week 5 vs. Week 10 vs. Week 17	High Physical Energy	Low Physical Energy	AVS_{r^*}
Mesor	0.47	0.41	0.44
Amplitude	0.36	0.39	0.38
Acrophase	0.15	0.15	0.15
$MAVS_R$			0.32

significant differences between groups within each time window, for each rhythm parameter and sensor feature. The Aggregate Variability Score (AVS) then aggregates these RVS values across all time windows for each rhythm parameter and sensor feature. Finally, the Mean AVS (MAVS) is calculated by averaging the AVS values across all rhythm parameters and all sensor features, respectively. Tables 7 and 8 (left-BGAT) present metrics comparing undergraduate and graduate students, as well as high and low energy levels across weeks 5, 10, and 17. In the context of our formulation, this corresponds to setting the comparison dimension $d \in D$ as a group combination (e.g., undergraduate vs. graduate), the evaluation context $e^* \in E$ as individual weeks, and computing $\mathrm{RVS}_f^{(D,e^*)}$, AVS_f, and MAVS_F as defined in Equations 1 and 2.

5.4.1 Rhythm Variability Between Undergraduate and Graduate Student Groups Across Weeks 5, 10, and 17.

Aggregating Across Rhythm Parameters. Table 7 (left-BGAT) highlights rhythm variability scores between undergraduate and graduate students across weeks 5, 10, and 17 for each individual rhythm parameter, across all sensor features. We observe that during the middle of the semester (week 10), the undergraduate and graduate student groups exhibit the greatest differences across all illustrated rhythm parameters (mesor 0.41, amplitude: 0.25, acrophase: 0.38), which suggests that academic and behavioral routines diverge most between these groups during mid-semester, potentially due to differing workloads, exam schedules, or adaptation to the academic environment. By the end of the semester, however, these groups' rhythms become highly aligned, as reflected by the lower RVS values (mesor: 0.04, amplitude: 0.08, acrophase: 0.21) across all sensor features. This convergence indicates that both undergraduate and graduate students experience similar end-of-semester behavioral patterns, such as reduced activity variability, synchronized sleep schedules, or consistent study habits, driven by shared academic responsibilities like final exams or deadlines.

The metrics aggregated across all weeks (AVS) show that mesor had the highest variability (0.26) over time between the two student groups, with acrophase being close (0.25), and amplitude (0.19) having the lowest

Table 8. Rhythm Variability Scores (RVS) are calculated between groups for each week (left-BGAT) and between weeks for each group (right-BTAG), for each target sensor feature s^* across all rhythm parameters. The Aggregate Variability Scores (AVS) then aggregate RVS values across all weeks (BGAT) and groups (BTAG) for a given sensor feature. Finally, Mean AVS (MAVS) is calculated by averaging AVS values across all sensor features. Higher RVS values indicate greater rhythmic divergence between groups in a given week (BGAT) and between weeks in a given group (BTAG), while higher AVS and MAVS values reflect greater overall variability across time (BGAT) and groups (BTAG) and sensor features, respectively. A few sample sensor features are shown, with zero values indicating that two varying dimensions are not significantly different from each other. A comprehensive version of this table with all sensor features is included in the appendix.

Between Groups Across Time Windows (BGAT) for All Rhythm Parameters				
Undergraduate vs. Graduate	Week 5	Week 10	Week 17	AVS_{s^*}
total duration on foot and on bicycle activities	0.32	0.99	0.98	0.76
standard deviation of steps	0.66	0.99	0.32	0.66
maximum steps	0.66	0.67	0.33	0.55
...	...	...	...	...
average battery consumption rate	0	0	0	0
$MAVS_S$				0.32
Low vs. High Cognitive Energy	Week 5	Week 10	Week 17	AVS_{s^*}
average heart rate	0.97	0.32	0	0.43
total duration of sedentary bouts	0.33	0.97	0	0.43
median heart rate	0.97	0.33	0	0.43
...	...	...	...	...
average battery consumption rate	0	0	0	0
$MAVS_S$				0.15
Low vs. High Emotional Energy	Week 5	Week 10	Week 17	AVS_{s^*}
average duration of sedentary bouts	0.33	0.65	0.33	0.43
time at home	0.33	0	0.65	0.33
maximum duration of sedentary bouts	0.32	0.65	0	0.32
...	...	...	...	...
average battery consumption rate	0	0	0	0
$MAVS_S$				0.12
Low vs. High Physical Energy	Week 5	Week 10	Week 17	AVS_{s^*}
total duration of sedentary bouts	0.65	0.66	0.32	0.55
maximum steps	0.65	0.33	0.65	0.55
maximum duration of sedentary bouts	0.66	0.66	0	0.44
...	...	...	...	...
average battery consumption rate	0	0	0	0
$MAVS_S$				0.17

Between Time Windows Across Groups (BTAG) for All Rhythm Parameters			
Week 5 vs. Week 10 vs. Week 17	Undergraduate	Graduate	AVS_{s^*}
time at home	0.90	0.24	0.57
average time spent in a cluster	0.75	0	0.37
minimum time spent in a cluster	0.75	0	0.37
...	...	...	...
average battery consumption rate	0	0	0
$MAVS_S$			0.19
Week 5 vs. Week 10 vs. Week 17	High Cognitive Energy	Low Cognitive Energy	AVS_{s^*}
time at home	0.66	0.82	0.74
total distance traveled	0.49	0.74	0.61
average time spent in a cluster	0.48	0.73	0.61
...	...	...	...
entropy heart rate	0	0	0
$MAVS_S$			0.29
Week 5 vs. Week 10 vs. Week 17	High Emotional Energy	Low Emotional Energy	AVS_{s^*}
time at home	0.83	0.65	0.74
minimum time spent in a cluster	0.65	0.72	0.69
average time spent in a cluster	0.65	0.64	0.65
...	...	...	...
entropy heart rate	0	0	0
$MAVS_S$			0.28
Week 5 vs. Week 10 vs. Week 17	High Physical Energy	Low Physical Energy	AVS_{s^*}
time at home	0.87	0.67	0.77
minimum time spent in a cluster	0.56	0.73	0.65
most common activity type	0.73	0.56	0.64
...	...	...	...
average battery consumption rate	0	0	0
$MAVS_S$			0.29

variability. This suggests that the groups differed most in their overall activity levels and timing of peak behaviors, while the strength or regularity of their rhythms remained more stable. The overall score across all rhythm parameters is 0.23, indicating differences in behavioral rhythms between undergraduate and graduate students throughout the semester, despite slight variability across specific parameters.

Aggregating Across Sensor Features. Table 8 (left-BGAT) highlights rhythm variability scores between undergraduate and graduate students across weeks 5, 10, and 17 for selected individual sensor features, across all rhythm parameters. We observe that, similar to the results in Table 7 and discussion in Section 5.4.1 regarding aggregating across rhythm parameters, during the middle of the semester (week 10), the undergraduate and graduate groups exhibit the greatest difference across all illustrated sensor features (total duration active: 0.99, std steps: 0.99, max steps: 0.67). These observations provide further evidence that academic and behavioral routines diverge most sharply mid-semester. In contrast, rhythm variability is lower during week 5 and lowest during week 17 (except for total duration active). This reinforces earlier insights, highlighting consistent patterns of temporal behavioral divergence between groups. Several sensor feature comparisons resulted in variances that are not statistically different, and are therefore reported as 0s. The scores aggregated across all weeks reveal substantial differences in activity-related behaviors between undergraduate and graduate students. Among the sensor features, total duration active exhibited the highest variability (0.76), followed by the standard deviation of steps (0.66) and maximum steps (0.55). These results suggest that the two groups differ meaningfully in their biobehavioral rhythms related to physical activity—potentially reflecting differences in academic schedules, campus engagement, or lifestyle patterns. The average variability score across all sensor features is 0.32, which is higher than the average observed across rhythm parameters. This indicates that the two student populations exhibit even greater divergence in observed behaviors than in modeled rhythms, underscoring the value of multimodal sensing for capturing nuanced group-level differences.

5.4.2 Rhythm Variability Between High and Low Energy Groups Across Weeks 5, 10, and 17.

Aggregating Across Rhythm Parameters. Table 7 (left-BGAT) highlights rhythm variability scores between high and low energy levels across weeks 5, 10, and 17 for each individual rhythm parameter, across all sensor features. We specifically examine differences across cognitive, emotional, and physical energy levels. When comparing high and low energy groups across rhythm parameters, we observe that week 17 consistently exhibits the lowest variability across most parameters and energy types, suggesting increasingly similar biobehavioral rhythms and a convergence in routines between groups. The exception is acrophase in the emotional (0.07) and physical energy (0.19) groups, where variability remains elevated, indicating persistent differences in the timing of peak behaviors despite overall convergence in other rhythm dimensions. These patterns echo those observed comparing graduate and undergradute student groups, with week 17 resulting in closer rhythmic patterns across groups.

The highest values are reported for amplitude, specifically for week 5, with values of 0.22 for cognitive energy groups, 0.18 for emotional energy groups, and 0.33 for physical energy groups. These early-semester differences in amplitude may reflect initial differences in motivation, physical activity routines, or lifestyle structure. The fact that amplitude, more than mesor or acrophase, shows the greatest early divergence also implies that energy level differences are most pronounced in the strength or intensity of behaviors, rather than their average levels or timing. However, these amplitude differences appear to diminish over time, indicating possible behavioral convergence between the two groups as the semester progresses.

Comparing aggregated scores across time, the physical energy groups exhibited the highest AVS values for each rhythm parameter, except for acrophase, relative to the cognitive and emotional energy groups. This suggests that individuals with high versus low physical energy demonstrated the greatest overall variation in their rhythmic patterns, particularly in mean levels (mesor) and strength of expression (amplitude). This trend is reinforced by the physical energy groups having the highest overall MAVS value (0.13), exceeding that of the cognitive and emotional groups, both of which had MAVS values of 0.11, indicating greater global variability in rhythms driven by physical energy level differences.

Aggregating Across Sensor Features. Table 8 (left-BGAT) highlights rhythm variability scores between low and high energy groups across weeks 5, 10, and 17 for selected individual sensor features, across all rhythm parameters. Selected sensor features significantly different between groups are shown. We specifically analyze differences across cognitive, emotional, and physical energy levels. We observe that across all energy-level groups, week 17 consistently exhibits the lowest rhythmic variance, indicating that by the end of the semester, high and low energy groups tend to converge in behavior. The exception is the time at home feature in the emotional energy group, which still shows some divergence. This supports earlier findings from rhythm parameter analyses, where week 17 also showed the highest similarity between groups. Among the energy groups, physical energy groups demonstrated the greatest variability, which aligns with the nature of the features analyzed (e.g., steps, sedentary bouts, physical activity). For the cognitive energy groups, heart rate-related features had high variability in week 5 (0.97), decreased in week 10 (0.33), and showed no variability in week 17—suggesting a convergence in physiological regulation over time. For emotional energy groups, sedentary bout features showed moderate variability in week 5 (0.32), peaked in week 10 (0.65), and dropped again in week 17 (0.33 for average duration, 0.00 for max duration), indicating that emotional energy may impact behavioral rhythms most during the mid-semester period. Finally, the physical energy groups maintained consistently high variability in sedentary and step-related features, with the strongest differences observed in week 5, slightly less in week 10, and the lowest in week

17—mirroring the general trend of behavioral convergence as the semester progresses. When aggregated over time, the individual sensor features have similar variabilities, ranging from 0.32-0.55 across the energy level groupings. The scores then averaged across all sensor features show that the variance between physical energy groups have the largest variability (0.17), followed by cognitive energy groups (0.15), and emotional energy groups (0.12).

5.5 Rhythm Variability Between Time Windows Across Groups (BTAG)

To compare rhythmic patterns between time windows across groups, we compute the Rhythm Variability Score (RVS), which quantifies rhythmic variance by measuring the proportion of rhythm parameters that show statistically significant differences between time windows, for each group, rhythm parameter, and sensor feature. The Aggregate Variability Score (AVS) then aggregates these RVS values across all groups for each rhythm parameter and sensor feature. Finally, the Mean AVS (MAVS) is calculated by averaging the AVS values across all rhythm parameters and all sensor features, respectively. Tables 7 and 8 (right-BTAG) highlight calculated metrics comparing between week 5, week 10, and week 17 across undergraduate and graduate students, as well as high and low energy levels.

5.5.1 Rhythm Variability Between Weeks 5, 10, and 17 Across Undergraduate and Graduate Student Groups.

Aggregating Across Rhythm Parameters. Table 7 (right-BTAG) highlights rhythm variability scores between weeks 5, 10, and 17 across undergraduate and graduate student groups for each individual rhythm parameter, across all sensor features. We observe that rhythm variability is substantially higher for undergraduate students compared to graduate students over time (e.g., mesor of 0.53 compared to 0.001). This suggests that undergraduate students experience greater fluctuations in the overall level, intensity, and timing of their behaviors on a week-to-week basis, whereas graduate students demonstrate more stable and consistent rhythms across the semester. The 0.001 mesor value for graduate students further highlights this consistency in their behavioral means over time. The aggregated scores across both the undergraduate and graduate groups reveal the highest variability for the mesor (0.27), followed by amplitude (0.23) and acrophase (0.15). This suggests that the mean levels vary the most, while timing of rhythms (acrophase) is relatively stable. Averaging across all rhythm parameters yields a MAVS value of 0.22, indicating a moderate level of rhythmic divergence between the two student populations over time.

Aggregating Across Sensor Features. Table 8 (right-BTAG) highlights rhythm variability scores between weeks 5, 10, and 17 across undergraduate and graduate student groups for each selected individual sensor feature, across all rhythm parameters. We observe that, consistent with our analysis aggregating across rhythm parameters, the undergraduate student group exhibits greater variability for each sensor feature shown compared to the graduate group. For undergraduates, the RVS for time at home reaches 0.90, while both average time spent in a cluster and minimum time spent in a cluster have variability scores of 0.75. In contrast, the graduate group shows no variability (0.00) for these features, except for time at home, which has a modest value of 0.24. These results reinforce the conclusion that undergraduate students display more behavioral fluctuation over time, whereas graduate students maintain more stable routines. When aggregating across groups, the AVS score is highest for time at home (0.57), highlighting it as the most variable behavioral feature between populations. Averaging across all sensor features yields a MAVS value of 0.19, indicating moderate overall divergence in behavioral rhythms between the two groups.

5.5.2 Rhythm Variability Between Weeks 5, 10, and 17 Across High and Low Energy Groups.

Aggregating Across Rhythm Parameters. Table 7 (right-BTAG) presents rhythm variability scores between weeks 5, 10, and 17 across high and low energy levels for each individual rhythm parameter, across all sensor features. We continue our analysis across cognitive, emotional, and physical energy levels. We observe that most variability scores are moderately high overall, indicating that substantial variability does exist between weeks across both high and low-energy groups. This suggests that weekly fluctuations in rhythmic patterns are present regardless of energy level. One exception to this pattern is acrophase. Acrophase consistently shows the lowest variability across all energy types, suggesting that the timing of peak behavior remained relatively stable over time, regardless of energy level. Interestingly, we observe contrasting patterns in acrophase variability across energy types over time. For the cognitive energy group, high-energy individuals exhibit substantially lower acrophase variability (0.09) compared to their low-energy counterparts (0.23), indicating more consistent timing in their behavioral rhythms. In contrast, within the emotional energy group, high-energy individuals display greater acrophase variability (0.22) than low-energy individuals (0.08), suggesting that emotional arousal may contribute to fluctuations in the timing of daily routines. These contrasting trends highlight that the impact of energy levels on rhythm stability is dimension-specific and may reflect underlying differences in cognitive control versus emotional response.

Aggregating the RVS values across groups reveals that mesor consistently exhibits the greatest variability, followed by amplitude, and then acrophase across all energy-level comparisons. This indicates that differences in the average level of behavior are most pronounced, while the timing of rhythmic peaks remains relatively stable. The overall MAVS value of 0.32 across group comparisons further supports this trend, reflecting a consistent and meaningful degree of rhythm variability across all energy dimensions.

Aggregating Across Sensor Features. Table 8 (right-BTAG) highlights rhythm variability scores between weeks 5, 10, and 17 across high and low energy groups for selected sensor features, across all rhythm parameters. The table highlights only those features that exhibited notable differences between energy groups. We continue our analysis by examining cognitive, emotional, and physical energy dimensions. Across all energy types, sensor-level variability remains consistently high, indicating that behavioral features fluctuate meaningfully over time between high and low energy groups. Within the cognitive energy groups, the low energy group exhibits consistently greater variability than the high energy group, particularly in time spent at home, total distance traveled, and average time spent in a cluster. This suggests that individuals with lower cognitive energy experience more fluctuation in their movement and location-based behaviors. For the emotional and physical energy groups, the patterns are more mixed. Interestingly, in contrast to the cognitive group, individuals with higher emotional or physical energy tend to show greater variability in home-stay behavior, suggesting more dynamic daily routines in these populations. Aggregating across groups, time at home emerges as the most variable feature overall, followed by minimum time spent in a cluster, particularly for emotional and physical energy comparisons. Finally, averaging variability scores across all sensor features yields a MAVS of 0.29 for both cognitive and physical energy groups, and 0.28 for the emotional energy group. These findings underscore that energy level distinctions, especially in cognitive and physical domains, manifest in consistent, measurable variations in behavioral rhythms over time.

Table 9. Rhythm variability scores (RVS) comparing two student groups for each week (left: BGAT) and between weeks for each group (right: BTAG) for each target rhythm parameter aggregated across all features. Metrics are computed as described in Table 7. Zero values indicate no significant rhythmic difference. AVS aggregates RVS across weeks, and MAVS averages AVS across parameters. Ours stands for our dataset. CE stands for the College Experience dataset.

Between Groups Across Time Windows (BGAT)								
Ours vs. CE	Nov 15	Nov 22	Nov 29	Dec 6	Dec 13	Dec 20	Dec 27	AVS_{r^*}
Mesor	0.665	0.667	0.999	0.991	0.667	0.666	0.937	0.799
Amplitude	0.666	0.661	0.973	0.976	0.645	0.655	0.653	0.747
Acrophase	0.320	0.000	0.955	0.965	0.664	0.332	0.660	0.557
$MAVS_R$								0.701

Between Time Windows Across Groups (BTAG)			
7 Weeks	Ours	CE	AVS_{r^*}
Mesor	0.299	0.171	0.235
Amplitude	0.350	0.116	0.233
Acrophase	0.119	0.000	0.060
$MAVS_R$			0.176

Table 10. Rhythm variability scores (RVS) comparing two student groups for each week (left: BGAT) and between weeks for each group (right: BTAG) for each feature across all rhythm parameters. Metrics are computed as described in Table 8. AVS_{s^*} aggregates RVS across weeks, and $MAVS_S$ averages AVS across features. Ours stands for our dataset. CE stands for the College Experience dataset.

Between Groups Across Time Windows (BGAT)								
Ours vs. CE	Nov 15	Nov 22	Nov 29	Dec 6	Dec 13	Dec 20	Dec 27	AVS_{s^*}
sum duration of unlock	0.985	0.667	0.991	0.995	0.999	0.999	0.999	0.948
sum of steps	0.667	0.664	0.986	0.988	0.656	0.661	0.323	0.706
duration in vehicle	0.000	0.000	0.999	0.993	0.332	0.000	0.979	0.472
$MAVS_S$								0.709

Between Time Windows Across Groups (BTAG)			
7 Weeks	Ours	CE	AVS_{s^*}
sum of steps	0.515	0.186	0.351
duration in vehicle	0.420	0.040	0.230
sum duration of unlock	0.092	0.005	0.049
$MAVS_S$			0.210

5.6 Additional Validation

To validate and highlight the value of our dataset and analytical method, we benchmark our approach with the College Experience dataset. We focus on the overlapping period between the two studies and analyze data from undergraduate students during the Fall 2021 semester. Due to differences in semester lengths and academic calendars at two universities, we select a seven-week window (2021-11-15 to 2022-01-02) that includes the final exam periods and partial winter breaks of both datasets. This period enables a comparison of rhythm variability between the two student groups and observes potential similarities and differences influenced by the university calendars and holiday events.

For this analysis, we include three sensor features shared across both datasets that had sufficient data: sum of steps, duration of screen unlock, and duration in vehicle. While the feature values may be influenced by external factors, such as campus environment and local weather conditions, they are sufficiently similar to demonstrate that our method supports cross-dataset comparisons.

Aggregating Across Sensor Features. Table 9 (BGAT, left) presents rhythm variability scores (RVS) between groups across time windows, calculated separately for each rhythm parameter. The results indicate strong between-group rhythm variability. Mesor and amplitude consistently exhibit high RVS values across all seven weeks, suggesting substantial and persistent differences in activity rhythms between the two student groups. For all three rhythm parameters, RVS values peak around November 29 and December 6, coinciding with differences in academic schedules. Students in the College Experience dataset were on break during these weeks, whereas students in our dataset were still in class. Beyond these two weeks, no simultaneous peaks are observed across all three parameters. In contrast to mesor and amplitude, acrophase RVS values remain relatively low outside the weeks of November 29 and December 6, as shown in its AVS_{r^*}. The value of 0.00 during the week of November 22 suggests that both groups exhibited similar timing of peak activity, likely due to the influence of Thanksgiving. Table 9 (BTAG, right) shows RVS values across weeks within each group for each of rhythm parameter. Our dataset shows higher within-group variability across all rhythm parameters compared to the College Experience dataset, reflecting greater temporal fluctuation in activity levels. Notably, acrophase RVS values in the College Experience dataset remain at 0.00, indicating stable daily peak activity times across the seven-week period.

Aggregating Across Rhythm Parameters. Table 10 (BGAT, left) shows that duration of screen unlock has the highest AVS_{s^*}, indicating the greatest between-group rhythm variability, followed by sum of steps and duration in vehicle. Meanwhile, RVS values for duration in vehicle are 0.00 during the weeks of November 15, November 22, and December 20, which are weeks likely influenced by shared events, such as Thanksgiving and Christmas. Additionally, we also observe an increase in RVS during the weeks of November 29 and December 6, when our university was in session, but the College Experience university was on break. Table 10 (BTAG, right) reveals greater inter-group rhythm variability in our dataset compared to the College Experience dataset for all features. While sum of steps and duration in vehicle exhibit high rhythm variability, duration of screen unlock remains relatively stable between weeks within each group, indicating that its rhythm pattern is stable within each student group between weeks but differs between the two groups in any given week.

Overall, we observe that the between-group variability is greater than the within-group variability, which reflects that the two student populations experienced different rhythm patterns while maintaining relatively stable patterns over time. Additionally, the alignment between peaks in RVS scores and the differences in academic calendars across two groups demonstrates our methodology's sensitivity to time-dependent changes. These observations highlight the value of our approach for comparing rhythm variability. Although the analysis is limited to three sensor features, it validates the capability of our method to compare rhythm patterns across different datasets.

6 DISCUSSION

Our novel dataset and methodology enable analysis of multidimensional cyclic behaviors within college students. By combining high-frequency passive sensing with daily self-reports collected over several months, our dataset offers a rich, longitudinal view into participants' biobehavioral rhythms and internal states. The combination of active and passive data facilitates a more holistic understanding of how daily routines and cyclic patterns manifest across diverse contexts. To harness the full potential of these data streams, we utilize existing methods in combination with a new technique designed specifically to extract, quantify, and compare rhythmic structures embedded in passively sensed behavioral signals.

We introduce and define variability metrics to examine the variance in cyclic behavior across multiple modalities and evaluation contexts. The first is Rhythm Variability Score (RVS), which quantifies the degree of rhythmic divergence between dimensions across an evaluation context for each fixed parameter. These values can then be aggregated across all evaluation contexts into an Aggregate Variability Score (AVS) to provide broader assessment of rhythmic divergence. These AVS values can be further averaged across all parameters to yield the mean AVS (MAVS). These hierarchical levels of analysis enable the capture of both localized differences and broader trends in rhythmic behavior. Future studies can leverage these metrics to quantitatively assess rhythm variability across diverse contextual dimensions, providing a scalable foundation for personalized and population-level behavioral analysis. We envision this framework improving the reliability of behavioral inferences in sensing platforms by accounting for rhythm stability and variability over time.

To validate the presence of rhythms in our dataset, we conducted periodicity analysis, confirming dominant 24-hour cycles across behavioral features, along with intra-day and multi-day rhythms. Significant periods followed two trends: some behaviors showed many rhythms at short periods across nearly all participants, while others clustered at intuitive period lengths (e.g., 8, 12, or 24 hours). These patterns varied by sensor and feature, highlighting distinct underlying rhythms and the complexity of human behavior across multiple timescales. We then analyzed the impact of external events on rhythms and found that daily rhythms remain stable during the fall and spring semesters but are significantly disrupted during winter break. Further, we used the Cosinor model to extract rhythm parameters that capture temporal dynamics beyond traditional features. For instance, revealing the consistency and periodicity of behaviors rather than just the overall behavior level, enabling more nuanced behavioral interpretation. These insights can inform the design of context-aware ubiquitous computing systems that adapt to users' natural rhythms, detect disruptions indicative of life changes or health issues, and deliver timely interventions or recommendations.

We applied our rhythm variability methods to our dataset to assess how rhythmic patterns vary across both group and temporal dimensions, evaluating changes over rhythm parameters and sensor features. We analyzed variances between groups across time windows, and also between time windows across groups. We specifically focused on comparing undergraduate and graduate students, as well as cognitive, emotional, and physical energy groupings, in addition to three key weeks in a university semester (week 5, week 10, and week 17).

Comparing rhythm variability between groups across time reveals that during the fall semester, *undergraduate and graduate students have the greatest rhythmic divergence mid-semester* across both sensor features and rhythm parameters. This is likely due to differing academic demands, however *their rhythms converge by the end of the term*, reflecting shared behavioral patterns around the end of the semester. Results additionally reveal substantial *differences between groups in activity-related behaviors*, indicating meaningful divergence in biobehavioral rhythms linked to physical activity, likely reflecting underlying lifestyle or routine differences. Our analysis also demonstrates differences in participants' rhythms based on their internal state each week. Comparing energy types across time, we see that *at the end of the semester, there is the lowest variability between energy groups*, suggesting similar biobehavioral rhythms and a convergence in routines between groups. *The highest variances are found in amplitude*, specifically for the beginning of the semester, between the different energy groupings. Early-semester differences in amplitude suggest that group disparities are most pronounced in behavioral intensity, likely due to motivation or routine differences, but then these differences diminish over time, indicating behavioral convergence. When aggregated over all time contexts, *physical energy groups exhibit the highest and most persistent variability*, especially in activity-related features like steps and sedentary bouts.

Comparing rhythm variability between weeks across group contexts reveals further insights. *Rhythm variability is substantially higher for undergraduate students compared to graduate students over time*, suggesting that undergraduate students experience greater fluctuations in the overall level, intensity, and timing of their behaviors on a week-to-week basis, whereas graduate students demonstrate more stable and consistent rhythms across the semester. This pattern holds across both rhythm parameters and sensor features, with *undergraduate students showing significantly greater variability in time spent at home and in clusters*, indicating that graduate students maintain more consistent location-based routines. Analyzing internal energy groups, the variability scores are generally high, indicating that weekly fluctuations in rhythmic patterns are present regardless of energy level. We observe that *the effect of energy levels on rhythm timing varies by energy type: high cognitive energy is linked to more stable daily timing, while high emotional energy corresponds to greater fluctuations*—suggesting dimension-specific influences of internal state on rhythm stability. We also find that *individuals with reported low cognitive energy show more variability in their movement and location behaviors*, while those with *high emotional or physical energy tend to have more dynamic time at home routines*, highlighting different patterns of behavioral fluctuation across energy types, with time at home being the most variable feature reported.

We further validated our method on the College Experience Dataset by comparing rhythm variability across a shared seven-week academic period. Using three common sensor features, our metrics revealed consistent differences in rhythm patterns—particularly in mesor and amplitude—between the two student groups. These differences aligned with academic calendar events, such as finals and winter break, supporting the ability of our rhythm divergence metrics to capture behavioral variation across time windows and groups.

Our analysis underscores the practical value of rhythm modeling for personalized behavioral insights. By revealing how different group categorizations exhibit distinct rhythmic patterns within the same dataset, they highlight the importance of contextual framing when interpreting behavioral variability. For the ubiquitous computing community, these insights can inform the design of adaptive, context-aware systems that tailor interventions to users' natural rhythms, detect disruptions indicative of health or life changes, and improve the timing and personalization of recommendations. Disruptions and irregularities in human rhythms have been found in patients with severe disorders (e.g., cancer [38], neurodegenerative diseases [60], and sleep disorders [34]), indicating they may offer novel opportunities for mobile health researchers to identify users at risk for such illnesses. Similarly, biobehavioral rhythms relate to humans' propensity to socialize [37, 66], indicating our data and methods may be useful for better understanding these cycles, as well as the external behaviors that relate to them. They also provide a framework for improving the reliability and interpretability of sensor-based behavioral inferences at both individual and population scales. In all, these tools will help passive monitoring systems to better understand how participants' behaviors change, and how their current data relates to the previously sensed information.

6.1 Limitations and Future Work

The focus of this paper is to introduce a longitudinal multimodal mobile dataset and corresponding methodology primarily designed for investigating cyclic biobehavior. While we provide an analysis of biobehavioral rhythm variation and stability across time windows and in population groups, we do not focus on a specific health outcome. However, our dataset contains sensor data and self-reports that can be used to study mental and physical health. The sensor data provides valuable insight into physical activity, physiology, and social behaviors, which can be connected to internal rhythms and external events. Change in biobehavioral rhythms, particularly circadian rhythms, is inherently associated with an individual's health and wellness. Similarly, the self-report data grants insight into participants' internal cognitive and emotional states, enabling future work to use our dataset to study the relationship between health indicators and the biobehavioral rhythms. Additionally, we collected numerous health and lifestyle questionnaires from each participant at the beginning and end of the study, which enable further analyses of the connection between behaviors and well-being.

Our dataset represents a subset of a single university population, which may limit the generalizability of our analysis to broader and more diverse populations. Factors such as age, occupation, cultural context, or geographic location may significantly influence behavioral rhythms. Our dataset consists of undergraduates and graduate students, making it slightly more diverse than existing passive sensing datasets, which only include undergraduates [42, 62]. However, including participants who are not university students will further improve the generalizability of the data. Our future plan includes expanding the study to collect data from different demographics to capture various lifestyles affecting biobehavioral rhythms.

Moreover, as is common with passive sensing studies, our dataset contains missing data. We addressed this through interpolation or, in cases of extensive data gaps, by excluding affected participants. While these methods help preserve dataset integrity, they may also introduce bias. Future research could explore more sophisticated imputation techniques or simulation-based approaches to better infer missing segments and assess their influence on rhythm analysis outcomes.

The dataset was collected towards the end of the COVID-19 pandemic. By September 2021, when onboarding began, few COVID-19 precautions remained in place at our university. Although strict measures had been enforced the previous year, the only remaining rule required masks to be worn indoors until March 21, 2022. After that date, masks were only required in classrooms and the university hospital. While these policies were in effect, they were minimally invasive and unlikely to significantly influence participant behavior. The timing of our data collection provides a unique opportunity to capture participants' behaviors and self-reports as they adjust to post-COVID college life.

Our Cosinor analysis successfully identified rhythmic patterns in the extracted behavioral features. However, these features reduce the original high-frequency data to an hourly resolution. This transformation is useful, as it converts raw sensor data into meaningful features, such as time spent at home. But, this may limit our ability to detect and analyze rhythm cycles shorter than one hour that are present in the finer-grained data. Similarly, we analyzed group rhythmic differences across weekly windows. This may obscure finer-grained, day-to-day variations in behavior. However, this approach offers several advantages: it captures broader daily rhythm patterns, accounts for weekday–weekend variability, minimizes the influence of transient noise, and increases robustness to missing data, thereby enabling more consistent and reliable modeling across participants. Future work could explore feature extraction at a more granular scope.

Another limitation involves using subjective energy ratings, reported daily by participants, to form energy-based groupings rather than longer multi-question questionnaires. Several questions in our daily surveys were adopted from the NASA TLX [29]. To ensure that participants did not feel burned out from completing the daily surveys, we designed it to be short and unobtrusive. There are no validated short questionnaires that collect information about emotional, cognitive, and physical energy. As such, we developed Likert scale questions that assessed each type of energy. Similar questions have been used in other datasets to evaluate participants' mood and emotions [9, 32, 65].

7 CONCLUSION

Despite the significance of behavioral cycles, research on passively sensed behavior has often overlooked cyclical rhythmic patterns. We present a longitudinal multimodal mobile dataset for capturing and modeling biobehavioral rhythms that includes high-frequency sensor data and daily self-reports from 166 participants collected over up to 16 months. Additionally, we introduce three new metrics for comparing biobehavioral rhythms among different groups and over time. We apply these techniques to our dataset, revealing distinct differences in rhythms between undergraduate and graduate students, as well as among participants with varying energy levels each week. We also observe clear variations in these rhythms between the academic semester and winter break. Our dataset and methods provide a strong foundation for future research on capturing and modeling human rhythms. For the ubiquitous computing community, our analysis insights can support context-aware systems that align with users' routines, detect disruptions linked to health or life changes, and improve the timing of interventions and recommendations.

Acknowledgment

This work was funded by the National Science Foundation (NSF-IIS-1816687).

References

[1] [n. d.]. Advanced Health Smartwatch | Fitbit Sense. https://www.fitbit.com/global/us/products/smartwatches/sense

[2] Saeed Abdullah, Mark Matthews, Elizabeth Murnane, Mashfiqui Rabbi, John A. Torres, Santosh Kumar, and Tanzeem Choudhury. 2016. Cognitive Rhythms: Unobtrusive and Continuous Sensing of Cognitive Awareness Using a Mobile Phone. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, USA, 178–189. doi:10.1145/2971648.2971712

[3] S. Abdullah, M. Matthews, E. L. Murnane, G. Gay, and T. Choudhury. 2014. Towards Circadian Computing. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 673–684. doi:10.1145/2632048.2632100

[4] S. Abdullah, E. L. Murnane, M. Matthews, and T. Choudhury. 2017. Circadian Computing: Sensing, Modeling, and Maintaining Biological Rhythms. In *Mobile Health*. Springer, 35–58. doi:10.1007/978-3-319-51394-2_3

[5] Nancy E. Adler, Elissa S. Epel, Grace Castellazzo, and Jeannette R. Ickovics. 2000. Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology* 19, 6 (Nov. 2000), 586–592. doi:10.1037/0278-6133.19.6.586

[6] Thomas W. Boonstra, Jennifer Nicholas, Quincy J. J. Wong, Felicity Shaw, Stephen Townsend, and Helen Christensen. 2018. Using Mobile Phone Sensor Technology for Mental Health Research: Integrated Analysis to Identify Hidden Challenges and Potential Solutions. *Journal of Medical Internet Research* 20, 7 (July 2018), e10131. doi:10.2196/10131

[7] Kirk Warren Brown and Richard M. Ryan. 2003. The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology* 84, 4 (2003), 822–848. doi:10.1037/0022-3514.84.4.822

[8] Sebastian Burchert, Anne Kerber, Johannes Zimmermann, and Christine Knaevelsrud. 2019. 14-day smartphone ambulatory assessment of depression symptoms and mood dynamics in a general population sample: comparison with the PHQ-9 depression screening. https://zenodo.org/record/3384860. Zenodo.

[9] Matteo Busso, Andrea Bontempelli, Leonardo Javier Malcotti, Lakmal Meegahapola, Peter Kun, Shyam Diwakar, Chaitanya Nutakki, Marcelo Rodas Britez, Hao Xu, Donglei Song, Salvador Ruiz-Correa, Andrea-Rebeca Mendoza-Lara, George Gaskell, Sally Stares, Miriam Bidoglia, Amarsanaa Ganbold, Altangerel Chagnaa, Luca Cernuzzi, Alethia Hume, Ronald Chenu-Abente, Roy Alia Asiku, Ivan Kayongo, Daniel Gatica-Perez, Amalia De Götzen, Ivano Bison, and Fausto Giunchiglia. 2025. DiversityOne: A Multi-Country Smartphone Sensor Dataset for Everyday Life Behavior Modeling. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* (2025).

[10] Daniel J Buysse, Charles F Reynolds III, Timothy H Monk, Susan R Berman, and David J Kupfer. 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research* 28, 2 (1989), 193–213.

[11] Michelle L Byrne, Monika N Lind, Sarah R Horn, Kathryn L Mills, Benjamin W Nelson, Melissa L Barnes, George M Slavich, and Nicholas B Allen. 2021. Using mobile sensing data to assess stress: Associations with perceived and lifetime stress, mental health, sleep, and inflammation. *DIGITAL HEALTH* 7 (Jan. 2021), 20552076211037227. doi:10.1177/20552076211037227 Publisher: SAGE Publications Ltd.

[12] Charles S Carver. [n. d.]. You Want to Measure Coping But Your Protocol's Too Long: Consider the Brief COPE. ([n. d.]).

[13] Matthew Clark and Afsaneh Doryab. 2024. Changing Your Tune: Lessons for Using Music to Encourage Physical Activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 2 (May 2024), 1–24. doi:10.1145/3659611

[14] Sheldon Cohen and Harry M. Hoberman. 1983. Positive Events and Social Supports as Buffers of Life Change Stress. *Journal of Applied Social Psychology* 13, 2 (1983), 99–125. doi:10.1111/j.1559-1816.1983.tb02325.x

[15] Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. 1983. A Global Measure of Perceived Stress. *Journal of Health and Social Behavior* 24, 4 (Dec. 1983), 385. doi:10.2307/2136404

[16] Sheldon Cohen, Robin Mermelstein, Tom Kamarck, and Harry M. Hoberman. 1985. Measuring the Functional Components of Social Support | SpringerLink. *Social Support: Theory, Research and Applications* 24 (1985). doi:10.1007/978-94-009-5115-0_5

[17] Sunny Consolvo, Ryan Libby, Ian Smith, James A. Landay, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, and Louis LeGrand. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press, Florence, Italy, 1797. doi:10.1145/1357054.1357335

[18] G. Cornelissen. 2014. Cosinor-based Rhythmometry. *Theoretical Biology and Medical Modelling* 11, 1 (April 2014). doi:10.1186/1742-4682-11-16

[19] Alia J. Crum, Peter Salovey, and Shawn Achor. 2013. Rethinking stress: The role of mindsets in determining the stress response. *Journal of Personality and Social Psychology* 104, 4 (April 2013), 716–733. doi:10.1037/a0031201

[20] Dana Diotte. [n. d.]. Sleep Data. https://www.kaggle.com/danagerous/sleep-data. Kaggle.

[21] Ensheng Dong, Hongru Du, and Lauren Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases* 20, 5 (2020), 533–534.

[22] A. Doryab, A. K. Dey, G. Kao, and C. A. Low. 2019. Modeling biobehavioral rhythms with passive sensing in the wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3 (2019), 1–21. Issue 1. doi:10.1145/3314395

[23] Afsaneh Doryab, Jun Ki Min, Jason Wiese, John Zimmerman, and Jason Hong. 2014. Detection of Behavior Change in People with Depression. (2014).

[24] Zachary Englhardt, Chengqian Ma, Margaret E. Morris, Chun-Cheng Chang, Xuhai "Orson" Xu, Lianhui Qin, Daniel McDuff, Xin Liu, Shwetak Patel, and Vikram Iyer. 2024. From Classification to Clinical Insights: Towards Analyzing and Reasoning About Mobile and Behavioral Health Data With Large Language Models. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)* 8, 2 (2024), 1–25. doi:10.1145/3659604

[25] Denzil Ferreira, Vassilis Kostakos, and Anind K. Dey. 2015. AWARE: Mobile Context Instrumentation Framework. *Frontiers in ICT* 2 (April 2015). doi:10.3389/fict.2015.00006

[26] Yusuke Fukazawa, Taku Ito, Tsukasa Okimura, Yuichi Yamashita, Takaki Maeda, and Jun Ota. 2019. Predicting anxiety state using smartphone-based passive sensing. *Journal of Biomedical Informatics* 93 (May 2019), 103151. doi:10.1016/j.jbi.2019.103151

[27] Beniamino Hadj-Amar, Bärbel Finkenstädt Rand, Mark Fiecas, Francis Lévi, and Robert Huckstepp. 2020. Bayesian model search for nonstationary periodic time series. *J. Amer. Statist. Assoc.* 115, 531 (2020), 1320–1335.

[28] Franz Halberg. 1969. Chronobiology. *Annual Review of Physiology* 31, 1 (1969), 675–726. doi:10.1146/annurev.ph.31.030169.003331

[29] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[30] T. H. Holmes and R. H. Rahe. 1967. The Social Readjustment Rating Scale. *Journal of Psychosomatic Research* 11, 2 (1967), 213–218. doi:10.1016/0022-3999(67)90010-4

[31] Gyuwon Jung, Sangjun Park, and Uichin Lee. 2024. DeepStress: Supporting Stressful Context Sensemaking in Personal Informatics Systems Using a Quasi-experimental Approach. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 18 pages.

[32] Soowon Kang, Woohyeok Choi, Cheul Young Park, Narae Cha, Auk Kim, Ahsan Habib Khandoker, Leontios Hadjileontiadis, Heepyung Kim, Yong Jeong, and Uichin Lee. 2023. *K-EmoPhone, A Mobile and Wearable Dataset with In-Situ Emotion, Stress, and Attention Labels.* doi:10.5281/zenodo.7606611

[33] Lin Sze Khoo, Mei Kuan Lim, Chun Yong Chong, and Roisin McNaney. 2024. Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches. *Sensors* 24, 2 (2024), 348.

[34] Jacqueline M. Lane, Jingyi Qian, Emmanuel Mignot, Susan Redline, Frank A. J. L. Scheer, and Richa Saxena. 2023. Genetics of circadian rhythms and sleep in human health and disease. *Nature Reviews Genetics* 24, 1 (Jan. 2023), 4–20. doi:10.1038/s41576-022-00519-z

[35] Mark Leary, Kristine Kelly, Catherine Cottrell, and Lisa Schreindorfer. 2013. Construct Validity of the Need to Belong Scale: Mapping the Nomological Network. *Journal of Personality Assessment* 95, 6 (2013), 610–624. doi:10.1080/00223891.2013.819511

[36] Mu Lin, Andrew T. Campbell, Tanzeem Choudhury, Nicholas D. Lane, Mashfiqui Mohammod, Xiaochao Yang, Hong Lu, Giuseppe Cardone, Shahid Ali, Afsaneh Doryab, and Ethan Berke. 2012. BeWell+: multi-dimensional wellbeing monitoring with community-guided user feedback and energy optimization. In *Proceedings of the conference on Wireless Health - WH '12*. ACM Press, San Diego, California, 1–8. doi:10.1145/2448096.2448106

[37] Ralph E. Mistlberger and Debra J. Skene. 2004. Social influences on mammalian circadian rhythms: animal and human studies. *Biological Reviews* 79, 3 (Aug. 2004), 533–556. doi:10.1017/S1464793103006353

[38] Maximilian Moser, Reiner Penter, Matthias Fruehwirth, and Thomas Kenner. 2006. Why Life Oscillates - Biological Rhythms and Health. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. 424–428. doi:10.1109/IEMBS.2006.259562 ISSN: 1557-170X.

[39] M. Moškon. 2020. CosinorPy: A Python Package for Cosinor-Based Rhythmometry. *BMC Bioinformatics* 21, 1 (October 2020). doi:10.1186/s12859-020-03830-w

[40] Tahsin Mullick, Ana Radovic, Sam Shaaban, and Afsaneh Doryab. 2022. Predicting Depression in Adolescents Using Mobile and Wearable Sensors: Multimodal Machine Learning–Based Exploratory Study. *JMIR Formative Research* 6, 6 (June 2022), e35807. doi:10.2196/35807 Company: JMIR Formative Research Distributor: JMIR Formative Research Institution: JMIR Formative Research Label: JMIR Formative Research Publisher: JMIR Publications Inc., Toronto, Canada.

[41] Elizabeth L. Murnane, Yekaterina S. Glazko, Jean Costa, Raymond Yao, Grace Zhao, Paula M. L. Moya, and James A. Landay. 2023. Narrative-Based Visual Feedback to Encourage Sustained Physical Activity: A Field Trial of the WhoIsZuki Mobile Health Platform. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (March 2023), 1–36. doi:10.1145/3580786 Publisher: Association for Computing Machinery (ACM).

[42] Subigya Nepal, Wenjun Liu, Arvind Pillai, Weichen Wang, Vlado Vojdanovski, Jeremy F. Huckins, Courtney Rogers, Meghan L. Meyer, and Andrew T. Campbell. 2024. Capturing the College Experience: A Four-Year Mobile Sensing Study of Mental Health, Resilience and Behavior of College Students during the Pandemic. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 38 (mar 2024), 37 pages. doi:10.1145/3643501

[43] Emma Pierson, Tim Althoff, and Jure Leskovec. 2017. Modeling Individual Cyclic Variation in Human Behavior. doi:10.48550/arXiv.1712.05748 arXiv:1712.05748 [cs.SI] Accepted at WWW 2018.

[44] Skyler Place, Danielle Blanch-Hartigan, Channah Rubin, Cristina Gorrostieta, Caroline Mead, John Kane, Brian P. Marx, Joshua Feast, Thilo Deckersbach, Alex "Sandy" Pentland, Andrew Nierenberg, and Ali Azarbayejani. 2017. Behavioral Indicators on a Mobile Sensing Platform Predict Clinically Validated Psychiatric Symptoms of Mood and Anxiety Disorders. *Journal of Medical Internet Research* 19, 3 (March 2017), e6678. doi:10.2196/jmir.6678 Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[45] Abhishek Pratap, Ava Homiar, Luke Waninger, Calvin Herd, Christine Suver, Josh Volponi, Joaquin Anguera, and Patricia Arean. 2022. Real-world behavioral dataset from two fully remote smartphone-based randomized clinical trials for depression. *Scientific Data* 9 (08 2022). doi:10.1038/s41597-022-01633-7

[46] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang, and Tanzeem Choudhury. 2015. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, Osaka Japan, 707–718. doi:10.1145/2750858.2805840

[47] Lenore Sawyer Radloff. 1977. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement* 1, 3 (1977), 395–401.

[48] Filip Raes, Elizabeth Pommier, Kristin D. Neff, and Dinska Van Gucht. 2011. Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology Psychotherapy* 18, 3 (2011), 250–255. doi:10.1002/cpp.702

[49] Roberto Refinetti, Germaine Cornélissen, and Franz Halberg. 2007. Procedures for numerical analysis of circadian rhythms. *Biological Rhythm Research* 38, 4 (2007), 275–325. doi:10.1080/09291010600903692

[50] Dan Russell, Letitia Anne Peplau, and Mary Lund Ferguson. 1978. Developing a Measure of Loneliness. Journal of Personality Assessment (1978).

[51] Sohrab Saeb, Emily G. Lattie, Konrad P. Kording, and David C. Mohr. 2017. Mobile Phone Detection of Semantic Location and Its Relationship to Depression and Anxiety. *JMIR mHealth and uHealth* 5, 8 (Aug. 2017), e7297. doi:10.2196/mhealth.7297 Company: JMIR mHealth and uHealth Distributor: JMIR mHealth and uHealth Institution: JMIR mHealth and uHealth Label: JMIR mHealth and uHealth Publisher: JMIR Publications Inc., Toronto, Canada.

[52] Gerald J. Schmeling. 1976. *The Cyclical Nature of Human Behavior*. Ph. D. Dissertation. ProQuest Dissertations and Theses. https://proxy1.library.virginia.edu/login?qurl=https%3A%2F%2Fwww.proquest.com%2Fdissertations-theses%2Fcyclical-nature-human-behavior%2Fdocview%2F302827506%2Fse-2%3Faccountid%3D14678 Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2023-02-18.

[53] Jane Shakespeare-Finch and Patricia L. Obst. 2011. The Development of the 2-Way Social Support Scale: A Measure of Giving and Receiving Emotional and Instrumental Support. *Journal of Personality Assessment* 93, 5 (Sept. 2011), 483–490. doi:10.1080/00223891.2011.594124
[54] Bruce W. Smith, Jeanne Dalen, Kathryn Wiggins, Erin Tooley, Paulette Christopher, and Jennifer Bernard. 2008. The brief resilience scale: Assessing the ability to bounce back. *International Journal of Behavioral Medicine* 15, 3 (Sept. 2008), 194–200. doi:10.1080/10705500802222972
[55] Christopher J. Soto and Oliver P. John. 2017. The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology* 113, 1 (July 2017), 117–143. doi:10.1037/pspp0000096
[56] Jing Xian Teo, Sonia Davila, Chengxi Yang, An An Hii, Chee Jian Pua, Jonathan Yap, Swee Yaw Tan, Anders Sahlén, Calvin Woon-Loong Chin, Bin Tean Teh, et al. 2019. Digital phenotyping by consumer wearables identifies sleep-associated markers of cardiovascular disease risk and biological aging. *Communications Biology* 2, 1 (2019), 361.
[57] Julio Vega, Meng Li, Kwesi Aguillera, Nikunj Goel, Echhit Joshi, Kirtiraj Khandekar, Krina C. Durica, Abhineeth R. Kunta, and Carissa A. Low. 2021. Reproducible Analysis Pipeline for Data Streams: Open-Source Software to Process Data Collected With Mobile Devices. *Frontiers in Digital Health* 3 (2021). doi:10.3389/fdgth.2021.769823
[58] Rui Wang, Felix J. Ebling, Santosh Kumar, and Tanzeem Choudhury. 2014. StudentLife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students Using Smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*. ACM, 3–14. doi:10.1145/2632048.2632054
[59] John Ware, Mark Kosinski, and Susan Keller. 1996. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care* 34, 3 (1996). https://journals.lww.com/lww-medicalcare/abstract/1996/03000/a_12_item_short_form_health_survey__construction.3.aspx
[60] Melissa Wilking, Mary Ndiaye, Hasan Mukhtar, and Nihal Ahmad. 2013. Circadian Rhythm Connections to Oxidative Stress: Implications for Human Health. *Antioxidants & Redox Signaling* 19, 2 (July 2013), 192–208. doi:10.1089/ars.2012.4889
[61] Xuhai Xu, Prerna Chikersal, Afsaneh Doryab, Daniella K. Villalba, Janine M. Dutcher, Michael J. Tumminia, Tim Althoff, Sheldon Cohen, Kasey G. Creswell, J. David Creswell, Jennifer Mankoff, and Anind K. Dey. 2019. Leveraging Routine Behavior and Contextually-Filtered Features for Depression Detection among College Students. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (Sept. 2019), 1–33. doi:10.1145/3351274
[62] Xuhai Xu, Xin Liu, Han Zhang, Weichen Wang, Subigya Nepal, Yasaman Sefidgar, Woosuk Seo, Kevin S. Kuehn, Jeremy F. Huckins, Margaret E. Morris, Paula S. Nurius, Eve A. Riskin, Shwetak Patel, Tim Althoff, Andrew Campbell, Anind K. Dey, and Jennifer Mankoff. 2023. GLOBEM: Cross-Dataset Generalization of Longitudinal Human Behavior Modeling. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 190 (jan 2023), 34 pages. doi:10.1145/3569485
[63] Runze Yan, Xinwen Liu, Janine Dutcher, Michael Tumminia, Daniella Villalba, Sheldon Cohen, David Creswell, Kasey Creswell, Jennifer Mankoff, Anind Dey, and Afsaneh Doryab. 2022. A Computational Framework for Modeling Biobehavioral Rhythms from Mobile and Wearable Data Streams. *ACM Trans. Intell. Syst. Technol.* 13, 3, Article 47 (March 2022), 27 pages. doi:10.1145/3510029
[64] Runze Yan, Xinwen Liu, Janine M. Dutcher, Michael J. Tumminia, Daniella Villalba, Sheldon Cohen, John D. Creswell, Kasey Creswell, Jennifer Mankoff, Anind K. Dey, and Afsaneh Doryab. 2024. Identifying Links Between Productivity and Biobehavioral Rhythms Modeled From Multimodal Sensor Streams: Exploratory Quantitative Study. *JMIR AI* 3, 1 (April 2024), e47194. doi:10.2196/47194
[65] Sofia Yfantidou, Christina Karagianni, Stefanos Efstathiou, Athena Vakali, Joao Palotti, Dimitrios Panteleimon Giakatos, Thomas Marchioro, Andrei Kazlouski, Elena Ferrari, and Šarūnas Girdzijauskas. 2022. *LifeSnaps: a 4-month multi-modal dataset capturing unobtrusive snapshots of our lives in the wild.* doi:10.5281/zenodo.7229547 Version 4.
[66] Cheng Zhang, Chee Wei Phang, Xiaohua Zeng, Ximeng Wang, Yunjie Xu, Yun Huang, and Noshir Contractor. 2015. Circadian Rhythms in Socializing Propensity. *PLOS ONE* 10, 9 (Sept. 2015), e0136325. doi:10.1371/journal.pone.0136325
[67] Tianyi Zhang, Miu Kojima, and Simon D'Alfonso. 2024. AWARE Narrator and the Utilization of Large Language Models to Extract Behavioral Insights from Smartphone Sensing Data. *arXiv preprint arXiv:2411.04691* (2024).
[68] Patrick Zippenfenig. 2023. *Open-Meteo.com Weather API.* doi:10.5281/zenodo.7970649

Appendix A: Comparison of Passively Sensed Behavior Datasets

Dataset	Sample Size	Duration	Continuous Data Collection	Daily Self-Reports	Physiology Sensing	Behavior Sensing	Sensors
Human Rhythms (Ours)	166	Up to 16 Months	✓	✓	✓	✓	Screen status, location, battery, calls, WiFi, activity recognition, steps, heart rate, activity level, sleep, floors
GLOBEM [62]	497	4 Years	✘	✘	✓	✓	Location, phone usage, calls, bluetooth, physical activity, sleep
College Experience [42]	215	4 Years	✓	✘	✘	✓	Sleep, conversations, physical activity, location, app usage, brain imaging data
Lifesnaps [65]	71	4 Months	✓	✓	✓	✘	Sleep, temperature, distance, exercise, heart rate mindfulness, stress
StudentLife [58]	48	10 Weeks	✓	✓	✘	✓	Sleep, conversation, physical activity, location, light, bluetooth, audio, WiFi, screen status, phone charge, app usage
DeepStress [31]	24	6 Weeks	✓	✓	✘	✓	location, physical activity, phone app usage
Diversity One [9]	782	4 Weeks	✓	✓	✘	✓	Connectivity, environment motion, position, app usage, device usage
Moodpath [8]	113	2 Weeks	✓	✓	✘	✓	-
Teo et al. [56]	482	Up to 11 Days	✓	✘	✓	✘	Sleep, physical activity, heart rate, Genomic DNA
K-EmoPhone [32]	77	7 Days	✓	✓	✓	✓	GPS, battery, calls, WiFi, connectivity, data traffic, ringer mode, screen, bluetooth, media entries, messages, calories, steps

Appendix B: Pre- and Post-Baseline Survey Questionnaires

Questionnaire	Description
Big Five Personality Index [55]	Models participants' personality along the dimensions of openness, conscientiousness, extroversion, agreeableness, and neuroticism. This specific version of the questionnaire also included 15 smaller personality facets.
The Status Ladder [5]	Participants rank, on a ten-point scale, how important they believe themseleves to be in their community.
Emotional, Physical and Cognitive Regulation	Collects participants' techniques to regulate sadness, tension, anger, and tiredness. They also report information about their sleep and exercise habits.
Brief COPE [12]	Identifies how participants regulate their stress.
Brief Resilience Scale [54]	Learns how participants recover from stressful experiences.
Social Media Usage	Participants report how frequently they check and post to social media.
12 Item Short Health Form [59]	Identifies the participants' health status and how their wellbeing interferes with daily functioning.
Cohen-Hoberman Inventory of Physical Symptoms (CHIPS) [14]	Reports how frequently participants have experiences different physical symptoms in the last two weeks.
Center for Epidemiological Studies Depression Scale (CES-D) [47]	Collects how frequently participants experience different symptoms of depression.
Pittsburgh Sleep Index (PSQI) [10]	Evaluates participants' sleep quality.
Perceived Stress Scale [15]	Determines how frequently participants experience common causes of stress.
College Student Stressful Event Checklist [30]	Asks whether participants have experienced various stressful events that may occurring during their time in college
Interpersonal Support Evaluation Checklist-12 (ISEL-12) [16]	Investigates how supported participants feel.
UCLA Loneliness Scale [50]	Evaluates how often participants experience symptoms of loneliness.
Need to Belong Scale [35]	Asks how important participants consider social connections and acceptance to be.
Mindful Attention Awareness Scale [7]	Identifies how conscious participants are of their activities and internal state during the day.
Self-Compassion Scale: Short Form [48]	Investigates how compassionate participants are to themselves.
Two-Way Social Support [53] Scale	Asks participants how they receive and provide emotional support
Educational Affective Forecasting	Participants report their expected academic performance for the semester, and how they will feel if they surpass or fail to meet their expectations.
3-Item Growth Mindset Measure	Evaluates whether participants believe practicing a skill is worth the required effort.
Stress Mindset Measure [19]	Asks participants whether stress is a harmful challenge or a growth opportunity
Drug and Alcohol Usage	Identifies how frequently participants engage in drinking and smoking.

Appendix C: Distribution of Days of Data

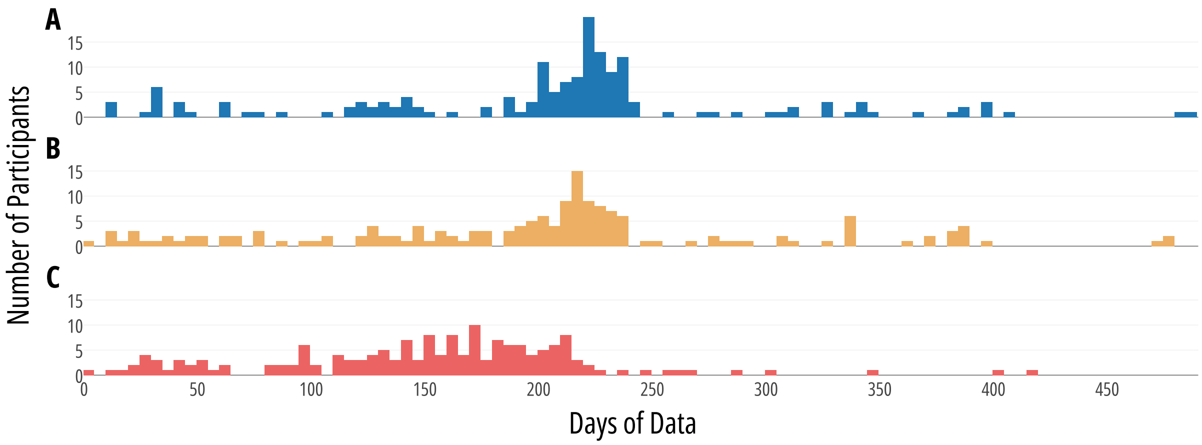


Fig. 4. The number of days of data we collected from participants for A) Fitbit, B) AWARE, C) Daily Surveys. We collected a median of 220 days of Fitbit data, 212 days of Aware data, and 160 days of survey data from each participant.

Appendix D: Mathematical Description

RVS

Equation 3 presents the RVS formulation for each target rhythm parameter, while Equation 4 defines the RVS for each target sensor feature. In both cases, the RVS is computed within a fixed evaluation context $e^* \in E$ by comparing across varying dimensions in D. A visual summary of calculations is shown in Table 11.

$$\mathrm{RVS}_{r^*}^{(D,e^*)} = \frac{1}{m}\sum_{a=1}^{m}\left(\frac{1}{C_D}\sum_{d\in D}\mathbf{1}\{\,p_{r^*s_a d}^{(e^*)} \le \alpha\,\}\left(1-\sum_{d\in D}p_{r^*s_a d}^{(e^*)}\mathbf{1}\{\,p_{r^*s_a d}^{(e^*)} \le \alpha\,\}\right)\right) \tag{3}$$

$$\mathrm{RVS}_{s^*}^{(D,e^*)} = \frac{1}{n}\sum_{a=1}^{n}\left(\frac{1}{C_D}\sum_{d\in D}\mathbf{1}\{p_{s^*r_a d}^{(e^*)} \le \alpha\}\left(1-\sum_{d\in D}p_{s^*r_a d}^{(e^*)}\,\mathbf{1}\{p_{s^*r_a d}^{(e^*)} \le \alpha\}\right)\right). \tag{4}$$

Table 11. Visual summary of Rhythm Variability Score (RVS) computed using derived Equations 3 and 4. An RVS is calculated for each fixed rhythm parameter r^* and each fixed sensor feature s^*, providing a RVS across pairwise comparison dimensions $d_{i,j} \in D$, where $i < j$, within a fixed evaluation context e^*.

		Rhythm Parameter $r^* = r_1$					Rhythm Parameter $r^* = r_n$				
		$d_{1,2}$	$d_{1,3}$	$\cdots$	$d_{\|D\|-1,\|D\|}$	$\mathrm{RVS}_{s^*}^{(D,h^*)}$	$d_{1,2}$	$d_{1,3}$	$\cdots$	$d_{\|D\|-1,\|D\|}$	$\mathrm{RVS}_{s^*}^{(D,e^*)}$
Sensor Feature $s^* =$	s_1	$p_{r_1 s_1 d_{1,2}}$	$p_{r_1 s_1 d_{1,3}}$	$\cdots$	$p_{r_1 s_1 d_{\|D\|-1,\|D\|}}$	$\mathrm{RVS}_{s_1}^{(D,e^*)}$	$p_{r_n s_1 d_{1,2}}$	$p_{r_n s_1 d_{1,3}}$	$\cdots$	$p_{r_n s_1 d_{\|D\|-1,\|D\|}}$	$\mathrm{RVS}_{s_1}^{(D,e^*)}$
	$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
	s_m	$p_{r_1 s_m d_{1,2}}$	$p_{r_1 s_m d_{1,3}}$	$\cdots$	$p_{r_1 s_m d_{\|D\|-1,\|D\|}}$	$\mathrm{RVS}_{s_m}^{(D,e^*)}$	$p_{r_n s_m d_{1,2}}$	$p_{r_n s_m d_{1,3}}$	$\cdots$	$p_{r_n s_m d_{\|D\|-1,\|D\|}}$	$\mathrm{RVS}_{s_m}^{(D,e^*)}$
$\mathrm{RVS}_{r^*}^{(D,e^*)}$		$\mathrm{RVS}_{r_1}^{(D,e^*)}$					$\mathrm{RVS}_{r_n}^{(D,e^*)}$				

AVS and MAVS

We can instantiate the general formulations of AVS and MAVS (Equation 2) for both rhythm-centric and sensor-centric cases. Specifically, we first compute the AVS for a fixed feature $f \in \{s^*, r^*\}$, and then average these scores across all rhythm parameters or sensor features to obtain MAVS, a global measure of rhythm difference. The transition from AVS to MAVS is denoted using a directional arrow. Table 12 illustrates the aggregation and averaging process.

$$\mathrm{AVS}_{r^*} = \frac{1}{|E|}\sum_{e\in E}\mathrm{RVS}_{r^*}^{(D,e)} \quad\Rightarrow\quad \mathrm{MAVS}_R = \frac{1}{n}\sum_{r\in R}\mathrm{AVS}_r \tag{5}$$

$$\mathrm{AVS}_{s^*} = \frac{1}{|E|}\sum_{e\in E}\mathrm{RVS}_{s^*}^{(D,e)} \quad\Rightarrow\quad \mathrm{MAVS}_S = \frac{1}{m}\sum_{s\in S}\mathrm{AVS}_s \tag{6}$$

Table 12. Visual representation of the computation of Mean Aggregate Variability Score (MAVS_F) from Aggregate Variability Scores (AVS_f). Each $\mathrm{RVS}_f^{(D,e)}$ is first aggregated across all comparison dimensions $d \in D$ for each evaluation context $e \in E$, then AVS_f is computed by averaging over all e. MAVS_F is obtained by averaging AVS_f over all $f \in F$.

Eval. Context $e \in E$	e_1	e_2	$\cdots$	$e_{\lvert E\rvert}$	AVS_f
$\mathrm{RVS}_{f_1}^{(D,e)}$	$\mathrm{RVS}_{f_1}^{(D,e_1)}$	$\mathrm{RVS}_{f_1}^{(D,e_2)}$	$\cdots$	$\mathrm{RVS}_{f_1}^{(D,e_{\lvert E\rvert})}$	AVS_{f_1}
$\mathrm{RVS}_{f_2}^{(D,e)}$	$\mathrm{RVS}_{f_2}^{(D,e_1)}$	$\mathrm{RVS}_{f_2}^{(D,e_2)}$	$\cdots$	$\mathrm{RVS}_{f_2}^{(D,e_{\lvert E\rvert})}$	AVS_{f_2}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\mathrm{RVS}_{f_{\lvert F\rvert}}^{(D,e)}$	$\mathrm{RVS}_{f_{\lvert F\rvert}}^{(D,e_1)}$	$\mathrm{RVS}_{f_{\lvert F\rvert}}^{(D,e_2)}$	$\cdots$	$\mathrm{RVS}_{f_{\lvert F\rvert}}^{(D,e_{\lvert E\rvert})}$	$\mathrm{AVS}_{f_{\lvert F\rvert}}$
MAVS_F					$\frac{1}{\lvert F\rvert}\sum_{f \in F}\mathrm{AVS}_f$

Appendix E: Variability Scores of Behavioral Features

Between Groups Across Time Windows (BGAT) for All Rhythm Parameters				
Undergradaute vs. Graduate	Week 5	Week 10	Week 17	AVS_{s^*}
total duration on foot and on bicycle activities	0.32	0.99	0.98	0.76
standard deviation of steps	0.67	1.0	0.33	0.66
total duration of vehicle activities	0.66	0.66	0.33	0.55
average number of steps	0.65	0.67	0.32	0.55
maximum steps	0.66	0.67	0.33	0.55
mean step cadence	0.65	0.67	0.32	0.55
minimum duration of sedentary bouts	0.64	0.33	0.66	0.55
maximum heart rate	0.66	0.67	0.32	0.55
average heart rate	0.66	0.66	0.32	0.55
sum of steps	0.65	0.67	0.32	0.55
mode heart rate	0.66	0.64	0.32	0.54
normalized location entropy	0.67	0.66	0.0	0.44
standard deviation of time spent in clusters	0.67	0.66	0.0	0.44
number of significant locations	0.67	0.66	0.0	0.44
number of location transitions	0.67	0.66	0.0	0.44
average movement speed	0.66	0.66	0.0	0.44
location entropy	0.67	0.66	0.0	0.44
average duration of sedentary bouts	0.32	0.33	0.65	0.44
median heart rate	0.66	0.66	0.0	0.44
maximum duration of sedentary bouts	0.32	0.65	0.33	0.43
standard deviation of heart rate	0.32	0.65	0.33	0.43
total duration of sedentary bouts	0.32	0.67	0.32	0.43
most common activity type	0.32	0.99	0.0	0.43
number of activity recognition activities	0.32	0.33	0.33	0.33
number of active wifi connections	0.66	0.0	0.33	0.33
difference between minimum and mode heart rate	0.0	0.66	0.33	0.33
variance of movement speeds	0.33	0.66	0.0	0.33
minimum heart rate	0.66	0.32	0.0	0.33
total distance traveled	0.33	0.66	0.0	0.33
difference between maximum and mode heart rate	0.32	0.65	0.0	0.32
time spent at location outliers	0.0	0.65	0.32	0.32
number of activity recognition events	0.0	0.33	0.33	0.22
number of sedentary bouts	0.0	0.33	0.32	0.22
time at home	0.0	0.0	0.66	0.22
average time spent in a cluster	0.32	0.32	0.0	0.22
minimum time spent in a cluster	0.33	0.33	0.0	0.22
maximum time spent in a cluster	0.0	0.32	0.0	0.11
number of device battery charge events	0.32	0.0	0.0	0.11
duration of phone discharges	0.32	0.0	0.0	0.11
ratio of stationary and moving time	0.0	0.33	0.0	0.11
number of device battery discharge events	0.0	0.33	0.0	0.11
duration of sedentary activities	0.0	0.0	0.0	0.0
average battery consumption rate	0.0	0.0	0.0	0.0
maximum battery consumption rate	0.0	0.0	0.0	0.0
duration of battery charge events	0.0	0.0	0.0	0.0
entropy of heart rate	0.0	0.0	0.0	0.0
duration with heart rate out of exercise zones	0.0	0.0	0.0	0.0
$MAVS_S$				0.32

Between Time Windows Across Groups (BTAG) for All Rhythm Parameters			
Week 5 vs. Week 10 vs. Week 17	Undergradaute	Graduate	AVS_{s^*}
time at home	0.9	0.24	0.57
average time spent in a cluster	0.75	0.0	0.37
minimum time spent in a cluster	0.75	0.0	0.37
most common activity type	0.75	0.0	0.37
mean step cadence	0.73	0.0	0.37
average number of steps	0.73	0.0	0.37
sum of steps	0.73	0.0	0.37
average movement speed	0.73	0.0	0.37
total duration of vehicle activities	0.5	0.16	0.33
total distance traveled	0.58	0.0	0.29
maximum steps	0.58	0.0	0.29
standard deviation of steps	0.58	0.0	0.29
location entropy	0.58	0.0	0.29
number of active wifi connections	0.5	0.08	0.29
time spent at location outliers	0.5	0.08	0.29
total duration on foot and on bicycle activities	0.58	0.0	0.29
standard deviation of time spent in clusters	0.58	0.0	0.29
number of significant locations	0.5	0.0	0.25
variance of movement speeds	0.5	0.0	0.25
number of location transitions	0.5	0.0	0.25
normalized location entropy	0.5	0.0	0.25
difference between maximum and mode heart rate	0.49	0.0	0.25
maximum duration of sedentary bouts	0.49	0.0	0.25
total duration of sedentary bouts	0.49	0.0	0.25
minimum duration of sedentary bouts	0.49	0.0	0.24
average duration of sedentary bouts	0.48	0.0	0.24
number of sedentary bouts	0.48	0.0	0.24
number of activity recognition activities	0.48	0.0	0.24
maximum time spent in a cluster	0.4	0.0	0.2
standard deviation of heart rate	0.33	0.0	0.16
difference between minimum and mode heart rate	0.24	0.0	0.12
duration of sedentary activities	0.23	0.0	0.12
number of activity recognition events	0.23	0.0	0.12
maximum heart rate	0.17	0.0	0.08
average heart rate	0.08	0.0	0.04
median heart rate	0.08	0.0	0.04
mode heart rate	0.08	0.0	0.04
duration of phone discharges	0.0	0.0	0.0
duration of battery charge events	0.0	0.0	0.0
maximum battery consumption rate	0.0	0.0	0.0
number of device battery discharge events	0.0	0.0	0.0
number of device battery charge events	0.0	0.0	0.0
ratio of stationary and moving time	0.0	0.0	0.0
average battery consumption rate	0.0	0.0	0.0
duration with heart rate out of exercise zones	0.0	0.0	0.0
minimum heart rate	0.0	0.0	0.0
entropy of heart rate	0.0	0.0	0.0
$MAVS_S$			0.19

Between Groups Across Time Windows (BGAT) for All Rhythm Parameters				
High vs. Low Cognitive Energy	Week 5	Week 10	Week 17	AVS$_{s^*}$
average heart rate	0.97	0.32	0.0	0.43
mode heart rate	0.97	0.33	0.0	0.43
median heart rate	0.97	0.33	0.0	0.43
total duration of sedentary bouts	0.33	0.97	0.0	0.43
mean step cadence	0.33	0.65	0.0	0.33
average number of steps	0.33	0.65	0.0	0.33
sum of steps	0.33	0.65	0.0	0.33
number of device battery charge events	0.32	0.0	0.65	0.32
number of location transitions	0.0	0.66	0.0	0.22
minimum heart rate	0.33	0.32	0.0	0.22
difference between minimum and mode heart rate	0.33	0.32	0.0	0.22
maximum steps	0.33	0.33	0.0	0.22
maximum duration of sedentary bouts	0.32	0.33	0.0	0.22
number of significant locations	0.0	0.66	0.0	0.22
number of activity recognition activities	0.33	0.32	0.0	0.22
time at home	0.33	0.33	0.0	0.22
maximum time spent in a cluster	0.0	0.65	0.0	0.22
standard deviation of steps	0.33	0.32	0.0	0.22
location entropy	0.0	0.66	0.0	0.22
number of sedentary bouts	0.32	0.32	0.0	0.21
number of activity recognition events	0.32	0.33	0.0	0.21
normalized location entropy	0.0	0.33	0.0	0.11
minimum time spent in a cluster	0.0	0.33	0.0	0.11
number of active wifi connections	0.0	0.33	0.0	0.11
average movement speed	0.32	0.0	0.0	0.11
variance of movement speeds	0.32	0.0	0.0	0.11
time spent at location outliers	0.0	0.32	0.0	0.11
total distance traveled	0.0	0.32	0.0	0.11
entropy of heart rate	0.32	0.0	0.0	0.11
average time spent in a cluster	0.0	0.33	0.0	0.11
duration of battery charge events	0.0	0.0	0.32	0.11
maximum heart rate	0.33	0.0	0.0	0.11
minimum duration of sedentary bouts	0.0	0.33	0.0	0.11
duration of sedentary activities	0.33	0.0	0.0	0.11
average duration of sedentary bouts	0.0	0.33	0.0	0.11
maximum battery consumption rate	0.0	0.0	0.0	0.0
most common activity type	0.0	0.0	0.0	0.0
average battery consumption rate	0.0	0.0	0.0	0.0
difference between maximum and mode heart rate	0.0	0.0	0.0	0.0
number of device battery discharge events	0.0	0.0	0.0	0.0
total duration on foot and on bicycle activities	0.0	0.0	0.0	0.0
duration of phone discharges	0.0	0.0	0.0	0.0
standard deviation of time spent in clusters	0.0	0.0	0.0	0.0
ratio of stationary and moving time	0.0	0.0	0.0	0.0
standard deviation of heart rate	0.0	0.0	0.0	0.0
duration with heart rate out of exercise zones	0.0	0.0	0.0	0.0
total duration of vehicle activities	0.0	0.0	0.0	0.0
MAVS$_S$				0.15

Between Time Windows Across Groups (BTAG) for All Rhythm Parameters			
Week 5 vs. Week 10 vs. Week 17	High Cognitive Energy	Low Cognitive Energy	AVS$_{s^*}$
time at home	0.66	0.82	0.74
total distance traveled	0.49	0.74	0.61
average time spent in a cluster	0.48	0.73	0.61
minimum time spent in a cluster	0.48	0.73	0.61
most common activity type	0.5	0.65	0.57
average movement speed	0.48	0.66	0.57
standard deviation of time spent in clusters	0.64	0.5	0.57
total duration of vehicle activities	0.5	0.58	0.54
standard deviation of steps	0.5	0.57	0.53
total duration on foot and on bicycle activities	0.57	0.49	0.53
normalized location entropy	0.5	0.5	0.5
location entropy	0.5	0.5	0.5
time spent at location outliers	0.5	0.5	0.5
variance of movement speeds	0.49	0.5	0.5
number of location transitions	0.49	0.5	0.5
sum of steps	0.5	0.49	0.49
average number of steps	0.5	0.49	0.49
mean step cadence	0.5	0.49	0.49
number of significant locations	0.49	0.5	0.49
number of active wifi connections	0.5	0.48	0.49
maximum steps	0.5	0.33	0.42
maximum time spent in a cluster	0.16	0.41	0.28
number of activity recognition activities	0.08	0.48	0.28
total duration of sedentary bouts	0.47	0.08	0.28
maximum heart rate	0.4	0.16	0.28
number of sedentary bouts	0.16	0.32	0.24
standard deviation of heart rate	0.32	0.08	0.2
maximum duration of sedentary bouts	0.25	0.16	0.2
number of activity recognition events	0.0	0.4	0.2
average heart rate	0.39	0.0	0.2
median heart rate	0.32	0.0	0.16
number of device battery charge events	0.24	0.08	0.16
difference between maximum and mode heart rate	0.16	0.16	0.16
average duration of sedentary bouts	0.25	0.0	0.12
minimum duration of sedentary bouts	0.24	0.0	0.12
duration of sedentary activities	0.16	0.08	0.12
minimum heart rate	0.24	0.0	0.12
mode heart rate	0.24	0.0	0.12
duration of battery charge events	0.24	0.0	0.12
maximum battery consumption rate	0.08	0.0	0.04
average battery consumption rate	0.08	0.0	0.04
duration of phone discharges	0.08	0.0	0.04
difference between minimum and mode heart rate	0.08	0.0	0.04
duration with heart rate out of exercise zones	0.08	0.0	0.04
number of device battery discharge events	0.0	0.08	0.04
ratio of stationary and moving time	0.0	0.0	0.0
entropy of heart rate	0.0	0.0	0.0
MAVS$_S$			0.29

Between Groups Across Time Windows (BGAT) for All Rhythm Parameters				
High vs. Low Emotional Energy	Week 5	Week 10	Week 17	AVS_{s^*}
average duration of sedentary bouts	0.33	0.65	0.33	0.43
time at home	0.33	0.0	0.65	0.33
minimum duration of sedentary bouts	0.32	0.33	0.32	0.32
maximum duration of sedentary bouts	0.32	0.65	0.0	0.32
number of device battery charge events	0.0	0.0	0.65	0.22
mean step cadence	0.0	0.33	0.32	0.22
total duration of sedentary bouts	0.0	0.65	0.0	0.22
number of sedentary bouts	0.33	0.0	0.32	0.22
sum of steps	0.0	0.33	0.32	0.22
average number of steps	0.0	0.33	0.32	0.22
difference between minimum and mode heart rate	0.33	0.33	0.0	0.22
standard deviation of steps	0.0	0.65	0.0	0.22
maximum steps	0.0	0.65	0.0	0.22
location entropy	0.66	0.0	0.0	0.22
number of location transitions	0.66	0.0	0.0	0.22
normalized location entropy	0.66	0.0	0.0	0.22
average movement speed	0.65	0.0	0.0	0.22
number of significant locations	0.66	0.0	0.0	0.22
entropy of heart rate	0.32	0.0	0.33	0.21
standard deviation of time spent in clusters	0.32	0.32	0.0	0.21
standard deviation of heart rate	0.33	0.0	0.0	0.11
median heart rate	0.32	0.0	0.0	0.11
mode heart rate	0.32	0.0	0.0	0.11
maximum heart rate	0.33	0.0	0.0	0.11
duration with heart rate out of exercise zones	0.0	0.0	0.33	0.11
duration of battery charge events	0.0	0.0	0.33	0.11
time spent at location outliers	0.32	0.0	0.0	0.11
number of active wifi connections	0.33	0.0	0.0	0.11
average heart rate	0.32	0.0	0.0	0.11
average time spent in a cluster	0.0	0.0	0.32	0.11
number of device battery discharge events	0.0	0.0	0.0	0.0
difference between maximum and mode heart rate	0.0	0.0	0.0	0.0
total distance traveled	0.0	0.0	0.0	0.0
duration of sedentary activities	0.0	0.0	0.0	0.0
number of activity recognition events	0.0	0.0	0.0	0.0
total duration of vehicle activities	0.0	0.0	0.0	0.0
most common activity type	0.0	0.0	0.0	0.0
total duration on foot and on bicycle activities	0.0	0.0	0.0	0.0
maximum battery consumption rate	0.0	0.0	0.0	0.0
average battery consumption rate	0.0	0.0	0.0	0.0
ratio of stationary and moving time	0.0	0.0	0.0	0.0
minimum time spent in a cluster	0.0	0.0	0.0	0.0
maximum time spent in a cluster	0.0	0.0	0.0	0.0
duration of phone discharges	0.0	0.0	0.0	0.0
number of activity recognition activities	0.0	0.0	0.0	0.0
minimum heart rate	0.0	0.0	0.0	0.0
variance of movement speeds	0.0	0.0	0.0	0.0
$MAVS_S$				0.12

Between Time Windows Across Groups (BTAG) for All Rhythm Parameters			
Week 5 vs. Week 10 vs. Week 17	High Emotional Energy	Low Emotional Energy	AVS_{s^*}
time at home	0.83	0.65	0.74
minimum time spent in a cluster	0.65	0.72	0.69
average time spent in a cluster	0.65	0.64	0.65
most common activity type	0.73	0.49	0.61
number of active wifi connections	0.72	0.48	0.6
average movement speed	0.66	0.5	0.58
location entropy	0.5	0.65	0.58
normalized location entropy	0.5	0.65	0.58
total distance traveled	0.66	0.5	0.58
number of significant locations	0.5	0.64	0.57
number of location transitions	0.57	0.57	0.57
total duration of vehicle activities	0.5	0.5	0.5
variance of movement speeds	0.5	0.5	0.5
time spent at location outliers	0.5	0.5	0.5
standard deviation of steps	0.5	0.5	0.5
standard deviation of time spent in clusters	0.41	0.58	0.5
maximum steps	0.58	0.33	0.46
total duration on foot and on bicycle activities	0.5	0.33	0.41
sum of steps	0.56	0.25	0.41
mean step cadence	0.56	0.25	0.41
average number of steps	0.56	0.25	0.41
maximum time spent in a cluster	0.31	0.32	0.32
total duration of sedentary bouts	0.33	0.24	0.28
number of sedentary bouts	0.24	0.31	0.27
median heart rate	0.47	0.0	0.23
mode heart rate	0.47	0.0	0.23
average heart rate	0.4	0.0	0.2
maximum duration of sedentary bouts	0.4	0.0	0.2
maximum heart rate	0.33	0.0	0.17
minimum duration of sedentary bouts	0.33	0.0	0.16
average duration of sedentary bouts	0.32	0.0	0.16
difference between minimum and mode heart rate	0.24	0.08	0.16
minimum heart rate	0.32	0.0	0.16
number of activity recognition activities	0.08	0.24	0.16
difference between maximum and mode heart rate	0.0	0.31	0.16
standard deviation of heart rate	0.08	0.22	0.15
duration of sedentary activities	0.0	0.15	0.08
duration with heart rate out of exercise zones	0.08	0.0	0.04
ratio of stationary and moving time	0.08	0.0	0.04
number of device battery discharge events	0.0	0.0	0.0
number of activity recognition events	0.0	0.0	0.0
average battery consumption rate	0.0	0.0	0.0
duration of phone discharges	0.0	0.0	0.0
duration of battery charge events	0.0	0.0	0.0
entropy of heart rate	0.0	0.0	0.0
maximum battery consumption rate	0.0	0.0	0.0
number of device battery charge events	0.0	0.0	0.0
$MAVS_S$			0.28

Between Groups Across Time Windows (BGAT) for All Rhythm Parameters				
High vs. Low Physical Energy	Week 5	Week 10	Week 17	AVS_{s^*}
maximum steps	0.66	0.33	0.65	0.55
total duration of sedentary bouts	0.65	0.67	0.33	0.55
sum of steps	0.33	0.66	0.32	0.44
mean step cadence	0.33	0.66	0.32	0.44
maximum duration of sedentary bouts	0.66	0.66	0.0	0.44
average number of steps	0.33	0.66	0.32	0.44
average duration of sedentary bouts	0.66	0.33	0.0	0.33
standard deviation of steps	0.33	0.33	0.33	0.33
minimum duration of sedentary bouts	0.66	0.33	0.0	0.33
average heart rate	0.65	0.0	0.33	0.33
mode heart rate	0.65	0.0	0.33	0.33
median heart rate	0.65	0.0	0.33	0.33
number of sedentary bouts	0.33	0.67	0.0	0.33
minimum heart rate	0.65	0.0	0.32	0.32
total distance traveled	0.66	0.0	0.0	0.22
average movement speed	0.66	0.0	0.0	0.22
time at home	0.66	0.0	0.0	0.22
variance of movement speeds	0.66	0.0	0.0	0.22
entropy of heart rate	0.32	0.0	0.33	0.21
duration of sedentary activities	0.33	0.0	0.0	0.11
standard deviation of heart rate	0.0	0.0	0.33	0.11
duration with heart rate out of exercise zones	0.0	0.0	0.32	0.11
maximum heart rate	0.0	0.0	0.33	0.11
difference between minimum and mode heart rate	0.0	0.0	0.32	0.11
difference between maximum and mode heart rate	0.0	0.0	0.33	0.11
number of significant locations	0.32	0.0	0.0	0.11
total duration on foot and on bicycle activities	0.0	0.32	0.0	0.11
number of location transitions	0.32	0.0	0.0	0.11
location entropy	0.32	0.0	0.0	0.11
number of activity recognition activities	0.33	0.0	0.0	0.11
average time spent in a cluster	0.32	0.0	0.0	0.11
number of active wifi connections	0.0	0.0	0.32	0.11
maximum battery consumption rate	0.0	0.0	0.0	0.0
time spent at location outliers	0.0	0.0	0.0	0.0
ratio of stationary and moving time	0.0	0.0	0.0	0.0
minimum time spent in a cluster	0.0	0.0	0.0	0.0
maximum time spent in a cluster	0.0	0.0	0.0	0.0
duration of phone discharges	0.0	0.0	0.0	0.0
duration of battery charge events	0.0	0.0	0.0	0.0
average battery consumption rate	0.0	0.0	0.0	0.0
number of device battery discharge events	0.0	0.0	0.0	0.0
number of device battery charge events	0.0	0.0	0.0	0.0
most common activity type	0.0	0.0	0.0	0.0
total duration of vehicle activities	0.0	0.0	0.0	0.0
normalized location entropy	0.0	0.0	0.0	0.0
number of activity recognition events	0.0	0.0	0.0	0.0
standard deviation of time spent in clusters	0.0	0.0	0.0	0.0
$MAVS_S$				0.17

Between Time Windows Across Groups (BTAG) for All Rhythm Parameters			
Week 5 vs. Week 10 vs. Week 17	High Physical Energy	Low Physical Energy	AVS_{s^*}
time at home	0.87	0.67	0.77
minimum time spent in a cluster	0.56	0.73	0.65
most common activity type	0.73	0.56	0.64
average time spent in a cluster	0.55	0.74	0.64
average movement speed	0.48	0.74	0.61
maximum steps	0.5	0.65	0.57
total distance traveled	0.48	0.66	0.57
standard deviation of steps	0.5	0.58	0.54
location entropy	0.5	0.58	0.54
variance of movement speeds	0.49	0.58	0.53
total duration on foot and on bicycle activities	0.73	0.33	0.53
total duration of vehicle activities	0.5	0.5	0.5
normalized location entropy	0.5	0.5	0.5
time spent at location outliers	0.5	0.5	0.5
number of significant locations	0.41	0.58	0.49
number of location transitions	0.49	0.5	0.49
number of active wifi connections	0.49	0.49	0.49
standard deviation of time spent in clusters	0.41	0.5	0.46
mean step cadence	0.49	0.33	0.41
sum of steps	0.49	0.33	0.41
average number of steps	0.49	0.33	0.41
maximum time spent in a cluster	0.16	0.47	0.32
total duration of sedentary bouts	0.41	0.16	0.29
number of sedentary bouts	0.41	0.16	0.29
difference between maximum and mode heart rate	0.0	0.49	0.24
maximum duration of sedentary bouts	0.41	0.0	0.2
maximum heart rate	0.32	0.08	0.2
duration of sedentary activities	0.08	0.24	0.16
median heart rate	0.31	0.0	0.16
mode heart rate	0.31	0.0	0.16
standard deviation of heart rate	0.23	0.08	0.16
number of activity recognition activities	0.0	0.31	0.15
minimum duration of sedentary bouts	0.25	0.0	0.12
average duration of sedentary bouts	0.25	0.0	0.12
average heart rate	0.24	0.0	0.12
duration with heart rate out of exercise zones	0.24	0.0	0.12
number of activity recognition events	0.0	0.24	0.12
entropy of heart rate	0.24	0.0	0.12
minimum heart rate	0.23	0.0	0.12
difference between minimum and mode heart rate	0.16	0.0	0.08
duration of battery charge events	0.08	0.0	0.04
number of device battery discharge events	0.0	0.0	0.0
maximum battery consumption rate	0.0	0.0	0.0
average battery consumption rate	0.0	0.0	0.0
number of device battery charge events	0.0	0.0	0.0
duration of phone discharges	0.0	0.0	0.0
ratio of stationary and moving time	0.0	0.0	0.0
$MAVS_S$			0.29