Triple Peak Day: Work Rhythms of Software Developers in Hybrid Work

Javier Hernandez, Vedant Das Swain, Jina Suh, Daniel McDuff, Judith Amores, Gonzalo Ramos, Kael Rowan, Brian Houck, Shamsi Iqbal, and Mary Czerwinski

Abstract—The future of work is rapidly changing, with remote and hybrid settings blurring the boundaries between professional and personal life. To understand how work rhythms vary across different work settings, we conducted a month-long study of 65 software developers, collecting anonymized computer activity data as well as daily ratings for perceived stress, productivity, and work setting. In addition to confirming the double-peak pattern of activity at 10:00 am and 2:00 pm observed in prior research, we observed a significant third peak around 9:00 pm. This third peak was associated with higher perceived productivity during remote days but increased stress during onsite and hybrid days, highlighting a nuanced interplay between work demands and work settings. Additionally, we found strong correlations between computer activity, productivity, and stress, including an inverted U-shaped relationship where productivity peaked at around six hours of computer activity before declining on more active days. These findings provide new insights into evolving work rhythms and highlight the impact of different work settings on productivity and stress.

Index Terms—Work-life balance, Productivity, Stress, Work Rhythms

I. INTRODUCTION

ORK rhythms, also referred to as the patterns and temporal cycles by which work activities are structured and carried out [1], have evolved significantly throughout history, reflecting changes in society and technology. In the past, work rhythms were often tied to the agricultural or industrial cycle, with set hours and breaks [2]. With the rise of the information age, however, work hours have become more flexible and variable, allowing for greater latitude in how workers organize their work [3]. More recently, the COVID-19 pandemic led to changes in work rhythms as numerous workers moved to remote work and adopted flexible work settings [4]. This transition has benefits in reducing commute and reclaiming more time for family responsibilities [5]-[7]. However, this blurring of boundaries can inadvertently lead workers to intensify work, lose their personal time, and, ultimately, find themselves without enough time to recuperate [6]– [9]. As work rhythms continue to evolve and change, it is imperative to better understand their potential impact on the well-being and productivity of workers.

With these changing work dynamics, the tools and platforms used to complete tasks have also evolved. Modern work,

J. Hernandez, Jina Suh, Judith Amores, Gonzalo Ramos, Kael Rowan, Brian Houck, Shamsi Iqbal, and M. Czerwinski are with Microsoft Research. Contact e-mail: javierh@microsoft.com

Vedant Das Swain is with Northeastern University Daniel McDuff is with Google Research particularly in the tech sector, relies heavily on computer usage. With software developers as a prime example, computer activity patterns can provide a lens into their daily rhythms, challenges, and behaviors. These digital footprints encode various signals - from levels of engagement [10], optimal productivity times [11], to moments of fatigue or burnout [12], [13]. Thus, monitoring and analyzing computer activity not only paints a picture of a day in the life of a developer but could be instrumental in optimizing work environments, work hours, and overall well-being in the age of flexible work rhythms. Through the lens of digital activity, this research contributes to a better understanding of work rhythms in software developers, and how they relate to important organizational variables such as work setting, stress, and productivity. In particular, this paper seeks to answer the following two research questions:

RQ1. How do work rhythms of computer activity vary across different work settings?

RQ2. How do different work rhythms relate to perceptions of productivity and stress levels?

To address these questions, we conducted an observational study in which we gathered anonymized computer activity data from 65 software developers over one month as they engaged in their regular work activities. Furthermore, we collected selfreports of stress and productivity, as well as their primary work setting for the day. While conventional methods for analyzing work rhythms focus on fluctuations in computer activity throughout the day, we introduce a third dimension that groups days with similar total daily activity volume. This enables us to capture not only the most common pattern of daily activity but also other less frequently observed rhythms that may vary based on the overall volume of work, which are then combined to represent what we call the "shape of work" (see Figure 1). Following this approach, our study uncovers significant differences across the different work settings, such as a strong positive correlation between late work and daily activity, as well as work setting. Additionally, we observed that the number of daily minutes of computer activity was positively correlated with stress and productivity, with a decrease in productivity on the most active days.

This work is organized as follows. In the next section, we review previous research on work rhythms and relevant studies focused on understanding the productivity and stress of software developers. We then describe our methodology including the data collection as well as the data analysis. We then

0000-0000/00\$00.0pr@s20011hteFresults to help answer each of the previous research

questions. We then discuss the findings and their potential implications in the context of individual and organizational well-being. Finally, we provide some conclusions and outline directions for future work.

II. RELATED WORK

A. Work Rhythms

Researchers have extensively studied work rhythms to better understand potential patterns in computer usage, collaboration, attention, and overall work engagement. While much of the focus has been on analyzing direct computer activity [1], [14], researchers have also explored other alternative signals, such as facial expressions [15] and physical activity [16], to gain a deeper understanding of user behavior and experience. In a seminal study, Begole et al. [1] visualized patterns of computer activity data from 18 information workers to help increase awareness, coordination, and communication among distributed groups. By displaying separate visualizations for each computer user, the authors demonstrated significant variance between individuals and within individuals throughout the day, as well as differences across work settings (office and home). Although no average was computed for all users, most of the visualizations showed that the majority of activity falls under two primary times of the day (9:00 am -12:00 pm and 1:00 pm-6:00 pm) which align well with traditional work hours. For one of the participants, researchers also showed a smaller third area of activity late at night (9:00 pm-10:00 pm).

Building upon this foundational work, subsequent studies explored specific aspects of digital work dynamics. For example, Mark et al. [14] logged the digital activity of 32 information workers over five days and combined it with experience sampling to better understand attention rhythms, showing that they fluctuated with context and time. By aggregating data from all participants, the researchers observed two distinctive peaks of computer activity: one at around 10:00 am and a second one at around 2:00 pm. As the software was installed on the participant's office computers, however, rhythms outside the office such as those during the night were not captured. In a separate study, McDuff et al. [15] monitored the webcam activity and facial activity of 117 information workers to better understand the patterns of emotion regulation. By averaging the total number of faces detected in front of the computer for every hour of the day, the researchers also revealed a consistent double peak pattern of activity. As in the previous study, the facial data logger was installed on the office machine, missing information beyond traditional work hours. More recently, Cao et al. [17] gathered telemetry data from a large sample of onsite and remote workers at Microsoft to help better understand multitasking behaviors during virtual meetings in times of COVID. As part of the analysis, the researchers aggregated daily email, file, and meeting activity for several months and further validated the double peak pattern as those found in pre-COVID studies. In contrast to the previous study, this work monitored telemetry around the clock and showed a continued but decreasing trend of computer activity at the end of the day. However, as these graphs aggregated patterns from a very large number of people into a single curve, some potentially emerging patterns may have been missed.

To extend previous work, this study similarly monitors workers around the clock beyond traditional work hours and focuses on hybrid work settings in which we visualize and analyze rhythms across fully onsite, fully remote, and hybrid days. In addition to including traditional 2D curve visualizations for comparison, we extend them on a third dimension that allows us to further break down the days according to the total computer activity.

B. Stress & Productivity

Understanding and measuring productivity and stress is an important area of research due to their key roles in employee well-being, work satisfaction, and more. Due to their subjective nature, self-reports remain the gold standard for their measurements [12], [18], but researchers are exploring more objective and passive methods, such as tracking computer activity and using wearable devices.

Over the years, software developer productivity and its perceived metrics have gained significant attention. For example, Storey et al. [19] proposed a bidirectional relationship between software developer job satisfaction and perceived productivity, highlighting various social and technical factors influencing this relationship. In a separate study, Meyer et al. [20] investigated developers' perceptions of productivity, revealing that developers feel most productive when completing sizable tasks without major interruptions. However, observational data showed a paradox where developers frequently switched tasks while still feeling productive, indicating that there may often be a mismatch between perceptions of behavior and actual observed behaviors. Workplace stress has also gained increased interest in recent years. While many definitions exist, stress can be broadly defined as the change in psychological and physical experiences occurring when external situations challenge or threaten someone [21], [22]. Importantly, high levels of stress may not always be associated with negative experiences [23]. In a relevant study, Kuutila et al. [24] addressed the issue of time pressure in software development, a factor that impacts both the development process and developers' stress levels.

When examining the relationship between stress and productivity, researchers have often considered the Yerkes-Dodson law [25] which suggests that there is an "inverted U" relationship between the two. However, this theory has faced controversy and produced conflicting findings [26], possibly due to a lack of standardized methodology for measuring them. In a relevant study, Bui et al. [27] conducted a study showing that higher stress levels were significantly correlated with decreased productivity, particularly affecting work satisfaction, which was also consistent with the findings of Muse et al. [28]. Similarly, Nan and Harter [29] investigated the impact of budget and schedule pressure on software development and found a significant U-shaped relationship. Their study indicated that both excessive and insufficient budget pressure led to a slowdown in the development cycle, increasing both time and effort. This parallels the concept of stress and inverse productivity in the Yerkes-Dodson law, suggesting that optimal pressure levels are necessary for peak performance.

This study combines objective data from computer activity and self-reported productivity and stress measurements to have a more comprehensive understanding of developer productivity, stress, and overall well-being in evolving hybrid work settings. Analysis of different patterns reveals an inverted U-shaped relationship between productivity and stress, indicating that excessive or insufficient stress may be negatively associated with productivity.

III. METHODS

To answer the research questions, we conducted a fourweek observational study in which we monitored the computer activity of several software developers and analyzed them in relation to their daily self-reports. This study received prior approval from our institutional review board and is described in the following sections.

A. Computer Activity

During the study, computer activity was captured by a custom-made data logger that recorded whether there was any interaction with a particular program (either through keyboard or mouse) as well as the name of the executable of the program¹. If the user had any interaction with a particular application, the whole minute was considered active. These measurements were aggregated at an hourly level. For any given hour, we obtained the number of minutes for which the participant interacted with different applications, ranging from 0 (i.e., no interaction) to 60 minutes (i.e., constant interaction). To facilitate the analysis, we computed overall computer activity (i.e., interaction with any application) as well as computer activity associated with different types of activities that are relevant in the context of software development. In particular, we follow the same categorization proposed by Meyer et al. [20] (see Table I) which automatically decomposes computer activity based on the titles of the application windows into development-related activity, such as coding and version control (VC), and other frequently observed activities, such as email and meetings. However, it is important to note that the absence of activity does not mean that the employee was not working, since they could have been performing a task outside the computer (e.g., reading an article, meeting with someone in person) or performing a task that was not tracked. More details about the collection of data can be found at Spencer et al. [30].

B. Daily Self-reports

At the end of each workday, participants were asked to reflect on different aspects of their day. In particular, participants self-reported their perceived productivity by answering the 7-point Likert question that we adapted from [14]: "Overall, how productive do you feel you were today?" with the following possible answers: 1 - extremely low, 2 - very low, 3 - somewhat low, 4 - moderately, 5 - somewhat high, 6 - very high, and 7 - extremely high. To help operationalize productivity,

TABLE I CATEGORIZATION OF APPLICATIONS BASED ON [20]

Category	Activity		
Development			
Code	Reading/editing/navigating code		
Debug	Debugging		
Review	Performing code reviews		
TestApp	Testing application outside IDE		
Version Control	Reading/accepting/submitting changes		
Other	Other related to development		
Email	Reading/writing emails		
Meetings	Meetings and calls		
Planning	Editing work items/tasks/todos;		
	creating/changing calendar entries		
ReadWriteDoc	Reading/editing documents and other artifacts		
	such as pictures		
Browsing	Internet browsing		
Other	Anything else, such as breaks or changing		
	music		

participants were instructed to think of productivity as a multidimensional concept that includes but is not limited to being able to accomplish everything that was planned, feeling efficient when performing the work, feeling satisfied with what was accomplished, being able to manage their time effectively, and being able to perform high-quality work. Similarly, participants self-reported their stress levels by answering the 5point Likert question, similar to ones used in previous studies of worker behaviors [12], [31], [32]: "Considering today's work, how would you rate your level of stress?" with the following possible answers: 1 - not at all, 2 - slightly stressed, 3 - moderately stressed, 4 - very stressed, and 5 - extremely stressed. For this question, participants were reminded that stress can be understood as the change in psychological and physical experiences that occurs when external situations challenge or threaten the individual (e.g., giving a presentation, having overly packed days, being unable to separate work and life demands). In addition, participants were reminded that having high stress at work may not be a negative experience. Finally, participants were also asked to answer "Where did you spend your working hours today?" with the following possible answers: remote (e.g., home), hybrid (sometime in the office and sometime at home), and onsite (e.g., office). It is important to note that the term "office" was broadly defined as somewhere on the corporate facilities which may include, but is not limited to, the office spaces and meeting rooms. The productivity and stress self-reported questions were kept the same as in prior work [12], [33] to help facilitate potential comparisons. Participants received daily reminders to complete a Qualtrics survey containing three questions, as well as the date for which they were reporting. The timing of the reminders was personalized for each participant based on the end-of-day schedules they provided during the onboarding process. However, the email encouraged participants to complete the survey whenever they considered their workday to be over. If participants forgot to complete the survey, they would receive a friendly reminder the following morning. To allow for flexibility, participants were able to select the date when answering the survey, which facilitated providing

¹The data logger did not collect any content or filenames in the applications to help protect the privacy of participants.

missing responses as well as updating previous responses if their initial answers did not accurately reflect their activity (e.g., changing work setting after submitting their responses).

C. Data Overview

A total of 65 US-based software developers were recruited from a large technology corporation who received a \$100 gift card after successful completion of the study. This group consisted of 25 women, 39 men, and 1 person who identified as non-binary/gender diverse during the summer of 2022. The majority of participants (48) were in the 18-35 age range, while 17 were between 36 and 55 years old. The developer group employed diverse software development methodologies, with Agile/Scrum being the predominant approach, though its implementation varied across teams. They were involved in creating applications and services for a range of platforms, catering to a broad spectrum of both internal and external customers. The corporation has a flexible work policy that provides employees with flexibility when defining their work hours and setting. In our studied population, the median starting and ending times were 9:00 am and 5:00 pm, respectively. The average self-report response rate in our sample was 13.77 responses each (SD = 3.88), with a decreasing trend in response rates observed over the course of the study. This corresponds to an overall compliance rate of 75% (SD = 17.94%), where full participation was defined as completing 20 responses throughout the study's duration. Additionally, approximately 36% of the responses were submitted one day after their intended dates, and participants did not update responses they had already submitted. This suggests that participants preferred submitting delayed responses over revising prior submissions multiple times. The data collected spanned 891 days of activity: 506 days were reported as remote (56.8%), 293 days as onsite (32.9%), and 92 days as hybrid (10.3%). However, 12 developers always reported being remote (18.46%), 6 of them always reported being onsite (9.23%), and the remaining 47 reported a combination of settings (72.3%). The average daily stress level was 2.72 with a standard deviation of 1.46, a minimum value of 1, and a maximum value of 5. Similarly, the average daily productivity level was 4.33 with a standard deviation of 2.01, a minimum value of 1, and a maximum value of 7.

D. Data Analysis

Computer activity across different days can vary significantly due to the unique personal and professional demands that may emerge [1]. Due to this variability, prior work has often aggregated activity across many days to help identify emerging rhythms [1], [14], [15], [34]. This approach takes advantage of the central limit theorem to help achieve more stable and reliable estimates of the population mean.

Aggregating computer activity across a group of days leads to a single curve that represents the average activity level for the sample data (e.g., different groups in Figure 2). However, this approach tends to collapse the days with different patterns. If we want to more precisely capture the multiple ways in which people may engage with work, we need to find a way

Algorithm 1 Generation of the shape of work curves

Require: Days with computer activity data D, where each day $d \in D$ has activity data $A_d(t)$ for t = 1 to 24 hours; Window size w = 150 minutes; Offset o = 20 minutes

Ensure: Curves associated with shape of work C 1: for each day $d \in D$ do Compute total daily activity $T_d = \sum_{t=1}^{24} A_d(t)$ 2: 3: end for 4: Initialize list of curves $C \leftarrow \emptyset$ 5: $T_{\min} \leftarrow \min\{T_d : d \in D\}$ 6: $T_{\max} \leftarrow \max\{T_d : d \in D\}$ 7: for $s = T_{\min}$ to $T_{\max} - w$ step o do Define window $W \leftarrow [s, s + w]$ $D_W \leftarrow \{d \in D : T_d \in W\}$ 10: if $D_W \neq \emptyset$ then $\begin{array}{l} \text{for } t=1 \text{ to } 24 \text{ do} \\ A(t) \leftarrow \frac{1}{|D_W|} \sum_{d \in D_W} A_d(t) \\ \text{end for} \end{array}$ 11: 12: 13: Append curve A to C14: end if 15: 16: **end for** 17: **return** C

to create and visualize multiple curves together. In our study, we observed that days with similar overall computer activity during the day tended to show similar hourly patterns, so we grouped days based on the total computer activity during the day. In particular, we first computed the total number of minutes of computer activity for each day (see center of Figure 1), and then iteratively computed and stacked the daily curves associated with different daily activity levels. To extract each curve, we aggregated days with daily activity levels that fall within the range of a sliding window of 150 minutes. This window is then iteratively slid with an offset of 20 minutes to extract curves associated with different daily activity levels. For instance, the dashed rectangle in Figure 1 shows the window that includes all days with daily activity levels between 400 and 550 minutes which are then averaged to generate a single curve. Note that this new curve has a different daily activity, which corresponds to 445 minutes in the example shown in the figure. By iteratively repeating the process through all possible windows, we are then able to generate the shape of work (see bottom-right graph). The window parameters were empirically selected to obtain a shape that minimized the appearance of high-frequency changes, such as those associated with potential outliers while capturing the dynamic range of different rhythms. Similarly, stress and productivity levels were extracted from each window by averaging the ratings associated with the corresponding days. After the windowing process, the average stress level per window was 2.74 with a standard deviation of 0.11, a minimum value of 2.56, and a maximum value of 2.93. Similarly, the average productivity level per window was 4.46 with a standard deviation of 0.26, a minimum value of 3.92, and a maximum value of 4.74. Algorithm 1 presents the pseudocode used to create the shape of work figures. To

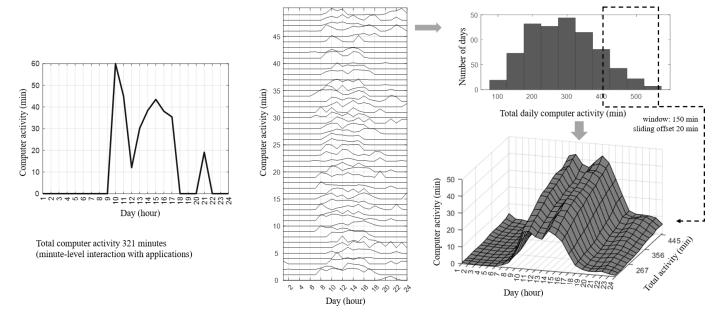


Fig. 1. Computation of the shape of work: 1) hourly computer activity is extracted for each day, 2) days with a similar number of minutes of computer activity are aggregated to create a curve, 3) each aggregation leads to a curve generating the shape of work.

facilitate the replicability of our methods, we provide scripts for computing curves and visualizing the shape of work in a GitHub repository².

To draw conclusions around the distribution of days across different settings, we created separate shape of work visualizations for all days (n: 891), remote (n: 506), onsite (n: 293), and hybrid (n: 92) settings, and color-coded them with the number of samples within each window. To study stress and productivity, we created the shape of work for all days (n: 891), and used different color mappings to represent the average self-reported ratings for stress and productivity. By considering this larger set of days, we were able to capture a broader range of ratings and better understand their relationship. For comparison purposes with prior work, we also included the traditional 2D curves in Figure 2 for the different groups of days as well as their \pm standard error.

To identify whether there are some significant differences in terms of computer activity over the day across different work settings (e.g., remote vs. hybrid vs. onsite), we performed a repeated measures ANOVA test and considered p <= 0.05 to be statistically significant. To compare specific metrics (e.g., overall computer activity) across multiple groups of days, we used the non-parametric Kruskal-Wallis test and similarly considered the p \leq 0.05 threshold to be statistically significant. To identify how different types of computer activity may contribute to stress and productivity, we used a linear mixed-effects model in which the dependent variable was either self-perceived stress or productivity, and the independent variables were the amount of daily activity observed for each of the application categories shown in Table I. To account for the varying frequencies of different activities (e.g., coding occurring more frequently than code reviews) and to facilitate a more meaningful comparison of their relative importance, each independent variable was standardized with z-score normalization. In addition, we introduced person and work setting as random variables to help account for their potential mediating role (see Equation 1).

$$Y \sim Dev._Code + Dev._Debug + Dev._Review + Dev._TestApp \\ + Dev._VC + Dev._Other + Email + Meetings + Planning \\ + ReadWriteDoc + Browsing + Other \\ + 1|Participant + 1|Work_Setting \\ Y \in \{stress, productivity\}$$

IV. RESULTS

This section describes the study findings in relation to our two research questions. To answer RQ1, we first characterize rhythms of computer activity (Section IV-A) and then evaluate how they change across work settings (Section IV-B). To answer RQ2, we first characterize the rhythms of perceived stress and productivity based on computer activity (Section IV-D and IV-C) and then examine how they relate to each other (Section IV-E).

A. What are the rhythms of computer activity?

Figure 2 shows the average pattern of computer activity on all days (red) as traditionally shown in previous work. As can be seen, we observe a very similar pattern to the one reported in recent studies [17] with two distinctive peaks at 10:00 am and 2:00 pm and a decreasing trend of activity over the day. To help capture a wider range of patterns, Figure 4 (top-left) shows the shape of work as proposed by this research. To help illustrate the number of days that fall under each window, each slice of the surface has been color-coded to indicate the number of aggregated days per window. In particular, dark blue and light yellow indicate the maximum (n: 418) and minimum (n: 105) number of days, respectively.

²https://github.com/microsoft/shape_of_work

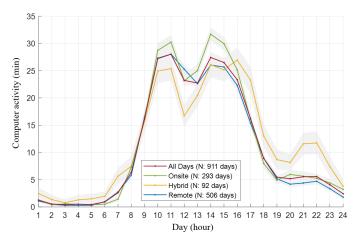


Fig. 2. Average computer activity pattern across different groups of days: all days (red), onsite (green), hybrid (yellow) and remote (blue). Grey areas indicate \pm standard error. Later in Figure 3 we expand this figure depiction of work with a 3D plot that we refer to as the "shape of work".

TABLE II Factors contributing to activity after 8:00 pm. (':p<0.1,*:p<0.05,**:p<0.01,***:p<0.001)

	Standardized Estimate	t-value	p-value
(Intercept)	2.051	2.251	0.025 *
Development Code	16.930	36.633	<0.001 ***
Development Debug	11.012	25.210	<0.001 ***
Development Review	0.407	0.892	0.373
Development TestApp	0.500	1.188	0.235
Development VC	0.549	1.307	0.192
Development Other	2.742	6.037	<0.001 ***
Email	12.566	22.178	<0.001 ***
Meetings	12.259	23.870	<0.001 ***
Planning	0.029	0.071	0.944
ReadWriteDoc	11.782	21.158	<0.001 ***
Browsing	27.276	57.990	<0.001 ***
Other	3.399	7.625	<0.001 ***

As can be seen, the curve with the largest number of samples (darker bluer color) has an average computer activity of around 289 minutes per day which captured around 47% of the days in our data (n: 418). Due to its large representation, this curve is the most closely aligned with the one shown in Figure 2 (red line). When examining the other colors, however, we observe that a significant number of days also show some work after 8:00 pm (48.60% of the days) which contributes to the emergence of a third peak of activity (around 9:00 pm). Furthermore, this third peak seems more pronounced on days with higher levels of computer activity during the day (>356 min/day) which cannot be captured with a 2D representation. In particular, work after 8:00 pm is significantly correlated with computer activity before 8:00 pm ($r_s(14) = 0.988$, p < 0.001).

To better understand the types of activities that may be performed after 8:00 pm, we performed a linear mixed-effects model like Equation 1 in which we used the amount of activity after 8:00 pm as a dependent variable. Table II shows the standardized estimates, t-values, and p-values for each of the estimates. As can be seen, a wide variety of activities seem to contribute to late work which includes Browsing, Dev. Code, Email, Meetings, ReadWriteDoc, Dev. Debug and Other.

B. How do the rhythms change across work settings?

Figure 2 shows the traditional average computer activity pattern when considering onsite (green), hybrid (yellow), and remote (blue). As can be seen, remote and onsite days show more similar shapes with relatively well-defined working hours between 8:00 am and 6:00 pm during which most of the activity falls. On the other hand, hybrid days tend to have more of a "rollercoaster" rhythm with ups and downs during the day and significantly more pronounced activity during the night (F(2,888) = 4.805, p = 0.008).

To better represent a wide range of patterns, Figure 3 shows the shape of work for each of the work settings. This visualization allows one to more easily inspect the percentage of days for different patterns according to the total computer activity. In particular, computer activity after 8:00 pm was significantly more likely to occur on hybrid and remote days (64.13% and 51.38%, respectively) vs. onsite days (39%). On average, the amount of activity after 8:00 pm was 43 min for hybrid, 25 min for onsite, and 18 min for remote days. In addition, we observe that hybrid days had an average activity level of 288 min/day, followed by onsite days with 274 minutes per day, and remote days 252 with minutes per day.

C. What are the rhythms of perceived stress?

Figure 4 (center) shows the shape of work with updated color mapping to indicate the average self-reported stress level at the end of each day. As can be seen, we observe a strong correlation between stress and computer activity, in which lower stress days (light yellow) tend to have lower computer activity and higher stress days (dark red) tend to have higher computer activity ($r_s(14) = 0.988$, p < 0.001). To better understand the contributors of stress, we performed a linear mixed-effects model in which we used the stress level as a dependent variable. As shown in Table III, we observed that the significant variables in descending order of their standardized estimates were: Dev. Code, Meetings, Read-WriteDoc, and Dev. Debug. When comparing the stress levels across different work settings, we did not find any statistical significance ($\chi^2(2) = 1.113$, p = 0.573). However, longer (with work after 8:00 pm) versus shorter days were significantly more stressful for onsite ($\chi^2(1) = 22.114$, p < 0.001) and close to significant for hybrid days ($\chi^2(1) = 3.340$, p = 0.068).

D. What are the rhythms of perceived productivity?

Figure 4 (right) shows the shape of work with updated color mapping to indicate the average productivity level at the end of each day. As in the case of self-reported stress, we observe a strong correlation between productivity and computer activity, in which lower productivity (light yellow) tends to have lower levels of computer activity and higher productivity (dark green) tends to be associated with higher levels of computer activity ($r_s(14) = 0.703$, p = 0.003). However, we observe a small decline in terms of productivity for the larger amounts of computer activity (light green slices) indicating that too much activity may start impacting productivity. To further understand the contributors of productivity,

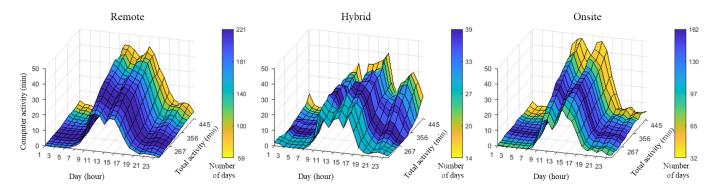


Fig. 3. Shape of work color-coded by the number of samples considered for each slice for remote (left), hybrid (center), and onsite (right) days.

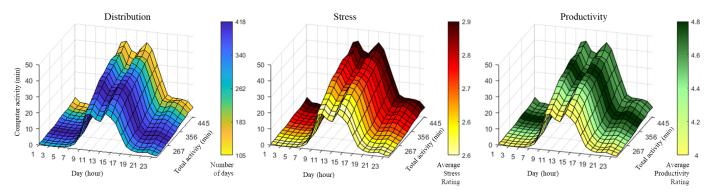


Fig. 4. Shape of work color-coded by the number of days considered for each slice (left), their average stress levels (center), and their average productivity levels (right).

TABLE III Factors contributing to end-of-day self-reported stress. ($\because p < 0.1, *: p < 0.05, **Y: p < 0.01, ***: p < 0.001)$

	Standardized Estimate	t-value	p-value
(Intercept)	2.082	19.652	<0.001 ***
Development Code	0.107	2.675	0.008 **
Development Debug	0.070	2.139	0.033 *
Development Review	0.021	0.635	0.525
Development TestApp	-0.015	-0.474	0.635
Development VC	0.022	0.543	0.587
Development Other	0.048	1.446	0.148
Email	-0.001	-0.032	0.975
Meetings	0.090	2.328	0.020 *
Planning	-0.040	-1.259	0.208
ReadWriteDoc	0.088	2.313	0.021 *
Browsing	-0.001	-0.035	0.972
Other	-0.003	-0.069	0.945

we similarly performed a linear mixed-effects model with productivity level as the dependent variable. As shown in Table IV, we observed that *Dev. Code* was the most significant contributing factor, followed by near significant estimates in *Browsing* and *Meetings*. When comparing the productivity levels across different work settings, we did not find any statistical significance ($\chi^2(2) = 3.456$, p = 0.178). However, longer (with work after 8:00 pm) vs. shorter days were rated as significantly more productive for remote days ($\chi^2(1) = 5.178$, p = 0.023).

	Standardized Estimate	t-value	p-value
(Intercept)	3.902	32.110	<0.001 ***
Development Code	0.232	4.552	<0.001 ***
Development Debug	0.004	0.099	0.921
Development Review	0.045	1.043	0.297
Development TestApp	-0.037	-0.902	0.368
Development VC	-0.054	-1.059	0.290
Development Other	-0.037	-0.836	0.403
Email	0.046	0.849	0.396
Meetings	0.089	1.788	0.074
Planning	0.002	0.057	0.955
ReadWriteDoc	0.033	0.681	0.496
Browsing	0.090	1.843	0.066
Other	0.002	0.039	0.969

E. How do computer activity, stress, and productivity relate to each other?

To analyze the potential relationship between stress and productivity in our study, Figure 5 jointly shows the average self-reported stress and productivity ratings for each of the slices considered when creating the shape of work. The figure also shows a cubic interpolation of the different points, the 95% prediction interval, and the associated average hours of activity per day. The coefficient of determination (R^2) was 0.95, suggesting that 95% of the variance can be accounted for by the model. Our data indicates that there is an inverted U-shape relationship between these two variables, indicating that an average computer activity of around 6 hours per day

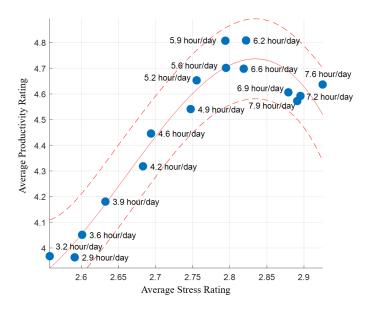


Fig. 5. Relationship between self-reported stress and productivity for multiple groups of days with a similar number of minutes of computer activity. The red continuous line shows a cubic interpolation and dashed red lines indicate the 95% prediction interval.

seems to yield optimal productivity.

V. DISCUSSION

Our study investigates the association between the shape of work in terms of computer activity and its association with work setting, stress, and productivity. In contrast to prior work [1], [14], [15], [17], our data reveals three distinctive peaks of activity at 10:00 am, 2:00 pm, and 9:00 pm. These peaks are separated by a decrease in computer activity which aligns with the potential occurrence of non-work activities such as having meals, commuting to work, or school-run schedules. However, these fluctuations could also be influenced by limitations in our data capture approach as discussed in the following section. By extending traditional work rhythms visualizations along a third dimension focused on the total volume of activity during the day, we also observed that late work (after 8:00 pm) was more likely to appear whenever there was a larger amount of activity during the day (before 8:00 pm). This finding supports that, on average, engaging in late work may serve as a buffer to be able to meet the increased demands of certain days as it could be generally expected.

Through the analysis of end-of-day self-reported data, we observed that stress levels were positively correlated with the total number of minutes of computer activity during the day and that perceptions of stress were associated with larger amounts of time spent on *Dev. Code, Meetings, ReadWriteDoc*, and *Dev. Debug*. These findings are consistent with a recent study by Morshed et al. [12] showing that work demands tend to be highly correlated with stress and computer activity. It is important to note, however, that different tasks may contribute to stress in different ways. For instance, high stress levels tend to occur whenever there is a high level of demands [35] which may be reflected in large amounts of coding activity for software developers. In this case, the stress associated with coding

would be related to the high cognitive demands involved and the associated increase in time spent on this activity. However, more infrequent tasks such as writing documents or having meetings can be perceived as inefficient or distracting [36] contributing to the overall stress level of developers. Similarly, our study also showed that productivity was also positively correlated with the number of daily minutes of computer activity. This positive correlation may be associated with spending more time at work and being able to achieve more, leading to higher perceived productivity. However, this effect only holds up to a certain point, after which additional time did not seem to contribute to productivity gains and could even have negative effects. Among the different activities, we found productivity to be associated with a larger amount of time spent on Dev. Code which is to be expected as it is required to achieve the main goals of software development [18], [37]. Furthermore, the ability to complete tasks has been shown [20] to be one key contributor to the perception of productivity in software developers.

Our study contributes to the understanding of the evolving nature of modern work, suggesting a nuanced interplay between work demands, work setting, and personal preferences. Specifically, our study indicates that the engagement and experience of late work may be moderated by the work setting during the day. On the one hand, our findings indicate that on remote workdays, longer days—those involving work during the so-called "third peak"—were more frequently associated with feelings of being productive compared to shorter remote days. While our study did not collect data to identify the contributing factors, this difference could indicate that late work helped individuals complete more tasks rather than merely helping them catch up due to other commitments (e.g., office commute, school-run schedules). On the other hand, our findings indicate that on onsite and hybrid workdays, longer days involving late work were more frequently associated with feelings of being stressed compared to shorter onsite/hybrid days. Similarly, this could be an indication that late work was more of a necessity to catch up with work rather than just completing more tasks. Our data also indicates that individuals engage in a wide variety of tasks during late work, indicating that late work may be used for different purposes probably depending on the unique professional needs of each individual [18]. For instance, participants working with teams located in different time zones may need to engage in meetings during the night [1]. In addition, some people may just prefer to engage in late work as there tend to be fewer interruptions. Overall, these findings further highlight the importance of providing agency to software developers [36], and how flexible job arrangements could help support the productivity and well-being of workers, especially when considering highly demanding days that may necessitate late work.

Stress and productivity are fundamental topics of research on organizational behavior which are challenging to study partly due to their subjective nature [28]. By combining computer activity levels with self-reports, we observed an "inverted U" relationship between stress and productivity, which highlights a delicate balance between the two that seems consistent with the Yerkes-Dodson law. Despite the

controversy of the law [26], this balance has been previously reported in a prior survey study focused on understanding the good days of software developers [36] and, to the best of our knowledge, this is the first time it has been observed in a real-life observational study of computer activity with software developers. In particular, we observe the existence of an optimal level of computer activity (around 6 hours) that maximizes productivity while maintaining moderate levels of stress in software developers. To perform this analysis, we considered all the data collected in the study irrespective of work setting, facilitating a broader range of productivity and stress ratings. However, even with a larger sample size, the range of average stress and productivity ratings (as shown in Figure 5) is still limited, especially when considering highstress levels. This limitation is due to the sampled population, which on average exhibited moderate stress levels, consistent with prior studies on information workers [13], [31], as well as the windowing process which effectively reduced the range of observed values. While the high coefficient of determination suggests a continuously decreasing trend in productivity with higher stress levels, future studies may consider more stressed populations to effectively cover a wider range of ratings, such as those experiencing burnout, which has been linked to significant workloads and other stressors [38]. Furthermore, collecting a larger number of days across different settings will help capture a broader range of productivity and stress ratings, facilitating more granular analysis and insights across work settings. In addition, our study analyzed computer activity at a higher level without incorporating intrinsic aspects of work tasks such as task difficulty or the worker's previous experience. However, prior work has shown that different factors may contribute to the specific shape of the curve. For instance, it has been hypothesized that lower-difficulty tasks may shift the curve to the right, indicating that higher stress levels may be more beneficial [26]. In the context of software developers, future work may also consider studying how well-known productivity factors such as the worker's previous experience on different tasks, work setup (e.g., laptop vs. desktop), the amount of interruptions [18], [33], or the use of AI chatbots [39] could influence this relationship.

VI. THREATS TO VALIDITY

This section highlights some of the limitations associated with the study, specifically regarding its construct, internal, and external validity.

A. Construct Validity

This study collected digital activity across different applications to understand the work rhythms of software developers. While this approach facilitates collecting objective data while minimizing the burden to participants [18], it is important to reiterate that computer activity does not capture all aspects of work. In a pre-COVID study by Meyer et al. [18], for instance, it was estimated that developers tend to spend about half of their work time away from computers. Furthermore, days with moderate computer activity could have been associated with

higher productivity due to increased time spent on creative brainstorming or other non-digital tasks that are less visible in telemetry data. It is also important to note that meetings during onsite days may be underrepresented in our data, as there may be ad hoc meetings not captured by digital platforms like MS Teams or Zoom. This could have reduced the representation of computer interactions during onsite work. In addition, our custom data logger did not collect interactions over the phone which may be more common in certain settings such as during the commute or late at night. To provide a more comprehensive analysis of work behavior, future efforts may consider measuring computer activity across multiple devices as well as incorporating experience sampling to help capture non-digital interactions (e.g., [12], [18]).

B. Internal Validity

This study combined computer activity with daily selfreports to capture subjective variables such as stress and productivity. Although self-reports are considered to be the gold standard for these metrics due to their large individual differences, it is important to note that self-reports are prone to well-studied recall biases (e.g., false memories, forgetting things, recency effect) [40]. Moreover, the awareness of being observed could lead to observation biases like the Hawthorne effect [41], influencing the way participants provide selfreports (e.g., overreporting productivity, underreporting stress). In addition, individual differences in personality traits and roles might also influence self-reports, as prior studies have shown that factors such as neuroticism and conscientiousness can shape how individuals perceive and report stress and productivity, respectively [42]. Future work could explore how the personality traits of developers, such as those captured by the Big Five personality model, and their specific roles (e.g., team lead, junior developer) impact correlations between stress, productivity, and computer activity, thus providing a more nuanced analysis. In our study, we opted to request self-reports at the end of the day and potential corrections in subsequent days to help minimize the potential burden on participants. However, future studies may consider other methods such as experience sampling during the day to more accurately capture the fluctuating nature of productivity and stress (e.g., [12], [14]). In addition, the incorporation of wearable sensors that track relevant physiological signals such as electrodermal activity and heart rate variability could help provide complementary objective insights (e.g., [12], [33]). Similarly, this study used self-reports to capture work settings but more objective approaches could have been used (e.g., tracking location through the smartphone GPS) to minimize errors and obtain a more granular representation of location during the day.

C. External Validity

The current study analyzed data from a sample of 65 software developers over one month. Focusing on software developers enabled a more specific examination of productivity and stress, but it is important to note that studying different professions might yield different results. For instance, the

optimal level of computer activity and contributing factors of stress and productivity may vary across information workers depending on the nature of their work. In addition, we recruited participants from the same company and during a specific time of the year which could have also influenced the patterns and findings. For instance, the number of developers on-call at a particular company can vary throughout the year, potentially leading to different activity patterns. In addition, self-reported stress levels might oscillate over the year, especially increasing during times of higher demand, such as before a product launch. Our particular study took place over the summertime which is usually associated with slightly higher than average activity than the rest of the year. Moreover, our study was conducted within a U.S.-based company, and cultural differences may impact work rhythms and behaviors. For example, prior research has shown that developers in Chinese IT companies tend to work longer hours compared to those in the United States, often due to different workplace cultures and expectations [43]. This suggests that our findings may not generalize to software developers in other cultural contexts, especially those with different norms regarding work hours and practices. Finally, all of our participants were under the age of 55, showing a potential bias towards a younger generation that encompasses individuals at various life stages, with diverse marital statuses and numbers of dependents. All these factors can potentially influence workers' capacity to work after 8 pm. In addition, circadian rhythms change as people age, leading to earlier wake times and potentially different work rhythms [44]. Future studies may consider more longitudinal experiments across a wide variety of companies, job roles, cultures, and age groups to help better understand the potential generalization of the findings.

VII. CONCLUSIONS

This study offers valuable insights into the work rhythms of software developers and how these are evolving in the context of hybrid work. In addition to the traditional doublepeak pattern, we identified a third peak of activity around 9:00 pm, which we then examined in relation to work setting, stress, and productivity. The triple-peak day underscores the complexities of hybrid work settings, where both work setting and daily activity can influence the perception of stress and productivity. Furthermore, we observed significant correlations between computer activity, stress, and productivity ratings, providing deeper insights into how these factors interact within the context of software development. We look forward to future studies that similarly broaden our understanding of work rhythms and their effects, contributing to the development of strategies and tools for a healthier and more productive future of work.

REFERENCES

- [1] J. B. Begole, J. C. Tang, R. B. Smith, and N. Yankelovich, "Work rhythms: analyzing visualizations of awareness histories of distributed groups," in *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, 2002, pp. 334–343.
- [2] J. Hassard, "Time and industrial sociology," in *Time, work and organization*. Routledge, 2017, pp. 13–34.

- [3] C. C. Kuhlthau, "The role of experience in the information search process of an early career information worker: Perceptions of uncertainty, complexity, construction, and sources," *Journal of the American Society for information Science*, vol. 50, no. 5, pp. 399–412, 1999.
- [4] D. Russo, P. H. Hanel, S. Altnickel, and N. Van Berkel, "The daily life of software engineers during the covid-19 pandemic," in 2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). IEEE, 2021, pp. 364– 373.
- [5] J. A. Mello, "Managing telework programs effectively," Employee responsibilities and rights journal, vol. 19, pp. 247–261, 2007.
- [6] S. M. Shaw, J. Andrey, and L. C. Johnson, "The struggle for life balance: Work, family, and leisure in the lives of women teleworkers," World Leisure Journal, vol. 45, no. 4, pp. 15–29, 2003.
- [7] D. Ford, M.-A. Storey, T. Zimmermann, C. Bird, S. Jaffe, C. Maddila, J. L. Butler, B. Houck, and N. Nagappan, "A tale of two cities: Software developers working from home during the covid-19 pandemic," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 31, no. 2, pp. 1–37, 2021.
- [8] C. Kelliher and D. Anderson, "Doing more with less? flexible working practices and the intensification of work," *Human relations*, vol. 63, no. 1, pp. 83–106, 2010.
- [9] C. A. Grant, L. M. Wallace, and P. C. Spurgeon, "An exploration of the psychological factors affecting remote e-worker's job effectiveness, well-being and work-life balance," *Employee Relations*, vol. 35, no. 5, pp. 527–546, 2013.
- [10] G. Mark, S. Iqbal, M. Czerwinski, and P. Johns, "Focused, aroused, but so distractible: Temporal perspectives on multitasking and communications," in *Proceedings of the 18th ACM Conference on Computer* Supported Cooperative Work & Social Computing, 2015, pp. 903–916.
- [11] H. Kaur, A. C. Williams, D. McDuff, M. Czerwinski, J. Teevan, and S. T. Iqbal, "Optimizing for happiness and productivity: Modeling opportune moments for transitions and breaks at work," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–15.
- [12] M. B. Morshed, J. Hernandez, D. McDuff, J. Suh, E. Howe, K. Rowan, M. Abdin, G. Ramos, T. Tran, and M. Czerwinski, "Advancing the understanding and measurement of workplace stress in remote information workers from passive sensors and behavioral data," in 2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE, 2022, pp. 1–8.
- [13] S. K. Nepal, J. Hernandez, J. Amores Fernandez, M. Bin Morshed, R. Lewis, H. Prafullchandra, and M. P. Czerwinski, "Workplace rhythm variability and emotional distress in information workers," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–8.
- [14] G. Mark, S. T. Iqbal, M. Czerwinski, and P. Johns, "Bored mondays and focused afternoons: the rhythm of attention and online activity in the workplace," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 3025–3034.
- [15] D. McDuff, E. Jun, K. Rowan, and M. Czerwinski, "Longitudinal observational evidence of the impact of emotion regulation strategies on affective expression," *IEEE Transactions on Affective Computing*, vol. 12, no. 3, pp. 636–647, 2019.
- [16] A. Doherty, D. Jackson, N. Hammerla, T. Plötz, P. Olivier, M. H. Granat, T. White, V. T. Van Hees, M. I. Trenell, C. G. Owen et al., "Large scale population assessment of physical activity using wrist worn accelerometers: the uk biobank study," PloS one, vol. 12, no. 2, p. e0169649, 2017.
- [17] H. Cao, C.-J. Lee, S. Iqbal, M. Czerwinski, P. N. Wong, S. Rintel, B. Hecht, J. Teevan, and L. Yang, "Large scale analysis of multitasking behavior during remote meetings," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–13.
- [18] A. N. Meyer, L. E. Barton, G. C. Murphy, T. Zimmermann, and T. Fritz, "The work life of developers: Activities, switches and perceived productivity," *IEEE Transactions on Software Engineering*, vol. 43, no. 12, pp. 1178–1193, 2017.
- [19] M.-A. Storey, T. Zimmermann, C. Bird, J. Czerwonka, B. Murphy, and E. Kalliamvakou, "Towards a theory of software developer job satisfaction and perceived productivity," *IEEE Transactions on Software Engineering*, vol. 47, no. 10, pp. 2125–2142, 2019.
- [20] A. N. Meyer, T. Fritz, G. C. Murphy, and T. Zimmermann, "Software developers' perceptions of productivity," in *Proceedings of the 22nd ACM SIGSOFT International Symposium on the Foundations of Software Engineering*. ACM - Association for Computing Machinery, November 2014.

- [21] R. S. Lazarus and S. Folkman, Stress, appraisal, and coping. Springer publishing company, 1984.
- [22] J. Blascovich and J. Tomaka, "The biopsychosocial model of arousal regulation," in *Advances in experimental social psychology*. Elsevier, 1996, vol. 28, pp. 1–51.
- [23] H. Selye, "Stress without distress," in Psychopathology of human adaptation. Springer, 1974, pp. 137–146.
- [24] M. Kuutila, M. V. Mantyla, U. Farooq, and M. Claes, "What do we know about time pressure in software development?" *IEEE Software*, vol. 38, no. 5, pp. 32–38, 2020.
- [25] R. M. Yerkes, J. D. Dodson *et al.*, "The relation of strength of stimulus to rapidity of habit-formation," 1908.
- [26] K. H. Teigen, "Yerkes-dodson: A law for all seasons," Theory & Psychology, vol. 4, no. 4, pp. 525–547, 1994.
- [27] T. Bui, R. Zackula, K. Dugan, and E. Ablah, "Workplace stress and productivity: a cross-sectional study," *Kansas journal of medicine*, vol. 14, p. 42, 2021.
- [28] L. A. Muse, S. G. Harris, and H. S. Feild, "Has the inverted-u theory of stress and job performance had a fair test?" *Human Performance*, vol. 16, no. 4, pp. 349–364, 2003.
- [29] N. Nan and D. E. Harter, "Impact of Budget and Schedule Pressure on Software Development Cycle Time and Effort," *IEEE Transactions on Software Engineering*, vol. 35, no. 5, pp. 624–637, 2009.
- [30] S. Buja, T. Zimmermann, P. Kirsanov, A. Tandon, X. Liu, J. Windsor, and H. Addis, "Understanding Software Developer Activity via Windows 10 Telemetry," Microsoft, Tech. Rep. MSR-TR-2019-42, April 2019, Presented at the 2019 Wharton People Analytics Conference. Finalist for the White Paper Competition.
- [31] E. Howe, J. Suh, M. Bin Morshed, D. McDuff, K. Rowan, J. Hernandez, M. I. Abdin, G. Ramos, T. Tran, and M. P. Czerwinski, "Design of digital workplace stress-reduction intervention systems: Effects of intervention type and timing," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–16.
- [32] V. Das Swain, V. Chen, S. Mishra, S. M. Mattingly, G. D. Abowd, and M. De Choudhury, "Semantic gap in predicting mental wellbeing through passive sensing," in *CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–16.
- [33] G. Mark, S. T. Iqbal, M. Czerwinski, P. Johns, A. Sano, and Y. Lutchyn, "Email duration, batching and self-interruption: Patterns of email use on productivity and stress," in *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2016, pp. 1717–1728.
- [34] M. Claes, M. V. Mäntylä, M. Kuutila, and B. Adams, "Do programmers work at night or during the weekend?" in *Proceedings of the 40th International Conference on Software Engineering*, 2018, pp. 705–715.
- [35] E. Demerouti, A. B. Bakker, F. Nachreiner, and W. B. Schaufeli, "The job demands-resources model of burnout." *Journal of Applied psychology*, vol. 86, no. 3, p. 499, 2001.
- [36] A. N. Meyer, E. T. Barr, C. Bird, and T. Zimmermann, "Today was a good day: The daily life of software developers," *IEEE Transactions on Software Engineering*, vol. 47, no. 5, pp. 863–880, 2019.
- [37] S. Astromskis, G. Bavota, A. Janes, B. Russo, and M. Di Penta, "Patterns of developers behaviour: A 1000-hour industrial study," *Journal of Systems and Software*, vol. 132, pp. 85–97, 2017.
- [38] S. Nepal, J. Hernandez, R. Lewis, A. Chaudhry, B. Houck, E. Knudsen, R. Rojas, B. Tankus, H. Prafullchandra, and M. Czerwinski, "Burnout in Cybersecurity Incident Responders: Exploring the Factors that Light the Fire," *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, no. CSCW1, pp. 1–35, April 2024.
- [39] OpenAI, "Gpt-4 technical report," 2023.
- [40] S. Shiffman, A. A. Stone, and M. R. Hufford, "Ecological momentary assessment annual review of clinical psychology, 4, 1-32," 2008.
- [41] A. L. Cohen, Hawthorne Effect. Cham: Springer International Publishing, 2023, pp. 33–35.
- [42] O. P. John and S. Srivastava, "The Big Five trait taxonomy: History, measurement, and theoretical perspectives," in *Handbook of personality: Theory and research*, Pervin, Lawrence A. and John, Oliver P., Eds. New York: Guilford Press, 1999, vol. 2, pp. 102–138.
- [43] J. Zhang, Y. Chen, Q. Gong, X. Wang, A. Y. Ding, Y. Xiao, and P. Hui, "Understanding the Working Time of Developers in IT Companies in China and the United States," *IEEE Software*, vol. 38, no. 2, pp. 96–106, 2021
- [44] T. H. Monk, "Aging human circadian rhythms: conventional wisdom may not always be right," *Journal of Biological Rhythms*, vol. 20, no. 4, pp. 366–374, 2005.