Continuous Alertness Assessments: Using EOG Glasses to Unobtrusively Monitor Fatigue Levels In-The-Wild

Benjamin Tag
Keio University
tagbenja@kmd.keio.ac.jp

Andrew W. Vargo
The Kyoto College of Graduate Studies for Informatics
vargo@kcg.ac.jp

Aman Gupta
Keio University
aman@kmd.keio.ac.jp

George Chernyshov
Keio University
chernyshov@kmd.keio.ac.jp

Kai Kunze
Keio University
kai@kmd.keio.ac.jp

Tilman Dingler
University of Melbourne
tilman.dingler@unimelb.edu.au

ABSTRACT

As the day progresses, cognitive functions are subject to fluctuations. While the circadian process results in diurnal peaks and drops, the homeostatic process manifests itself in a steady decline of alertness across the day. Awareness of these changes allows the design of proactive recommender and warning systems, which encourage demanding tasks during periods of high alertness and flag accident-prone activities in low alertness states. In contrast to conventional alertness assessments, which are often limited to lab conditions, bulky hardware, or interruptive self-assessments, we base our approach on eye blink frequency data known to directly relate to fatigue levels. Using electrooculography sensors integrated into regular glasses' frames, we recorded the eye movements of 16 participants over the course of two weeks in-the-wild and built a robust model of diurnal alertness changes. Our proposed method allows for unobtrusive and continuous monitoring of alertness levels throughout the day.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; *Human computer interaction (HCI)*; • **Hardware** → *Sensors and actuators*;

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300694

KEYWORDS

cognition-aware systems, circadian computing, fatigue, eye blink, electrooculography

ACM Reference Format:
Benjamin Tag, Andrew W. Vargo, Aman Gupta, George Chernyshov, Kai Kunze, and Tilman Dingler. 2019. Continuous Alertness Assessments: Using EOG Glasses to Unobtrusively Monitor Fatigue Levels In-The-Wild. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300694

1 INTRODUCTION

Human cognitive and physical performance are heavily dependent on the daily 24-hour cycle. While a biological rhythm, which is chronically out of sync, can cause serious health problems [18, 65], time-of-day variations have a significant impact on our everyday cognitive performance [56] affecting alertness and fatigue levels. This is due to the so-called *Homeostatic Process* (HP), which constitutes the increasing urge to sleep with prolonged wakefulness [11]. When working long hours this has been shown to lead to an increased risk of making mistakes and subsequently causing accidents [32]. Such hours are common practice in some professions, including pilots and medical staff, which demand extended work shifts [44]. On top of this, vehicles are often operated at late hours after a full day of work, where high fatigue levels and sleepiness has been shown to lead to delayed breaking reflexes [8, 36] with often fatal consequences.

Sleep deprivation leads to slower Reaction Times (RTs), which negatively affects task performance [57], including cognitive performance, which shows in a deterioration of vigilance and alertness levels [17, 66]. Vital biological signals, such as body temperature and heart rate, also underly the influence of our "biological clock" [45], which describes the endogenous, idiosyncratic fluctuations in wakefulness. These fluctuations are in part due to the *Circadian Rhythm* (CR), which is among others responsible for the post-lunch dip in vigilance despite a night of proper rest [2, 66].

Changes in alertness and fatigue also affect higher cognitive functions, such as reasoning and working memory [56]. It is, therefore, necessary to find ways to identify these changes in order for automated systems to be able to detect and predict these variations. The resulting cognition-aware systems are capable of identifying cognitive capacities and can dynamically adjust and respond to desirable and undesirable states, *e.g.*, by scheduling tasks effectively, triggering reminders to take a break in times of sleepiness, or turning off notifications to prevent interruptions when productivity is high [25, 52].

Traditional methods to assess alertness fluctuations include constrained settings or unpleasant and laborious procedures, such as extended measures in enclosed and controlled environments, so-called sleep labs, or repeated measurements of core body temperature through rectal probes [35, 45]. Recent developments in mobile and sensing technology have enabled everyday devices, such as smartphones, watches, and wristbands to become sophisticated trackers of our daily activities. They enable us to measure physiological signals, including heart-rate and blink rate around the clock. Data collected in this way has been shown feasible for analysis, detection, and monitoring of cognitive states: Abdullah *et al.* [3] and Dingler *et al.* [26] have proposed mobile solutions for tracking cognitive capacities (*e.g.*, alertness) based on data from smartphones. A major shortcoming of these proposed systems, however, is that the alertness measures are limited to phone usage, but fail to collect measures when the phone is not being used, such as while driving, in social gatherings, or during intense work sessions. Furthermore, since interactions with the smartphone or any other sensing device require active engagement, attention would be drawn away from the actual activity users were engaged with. To avoid these distractions and fill the occurring gaps of monitoring, we propose an unobtrusive, continuous logging technique based on monitoring eye movements.

To enable people to accurately track their fatigue levels in their everyday lives, we propose a solution utilizing sensing glasses to record Electrooculography (EOG) signals for detecting the characteristic eye movements occurring during eye blinks. Different studies have demonstrated that fatigue is directly related to changes in eye blink features, such as frequency and duration: greater fatigue causes higher Blink Frequency (BF) and longer blinks [12, 58]. In this research, we use an off-the-shelf eye-tracking device to unobtrusive collect eye movement data and detect changes of fatigue. We conducted an in-the-wild study to validate BF as a predictor of changing fatigue levels in everyday situations: for two weeks, participants periodically completed self-assessments in the form of psychomotor-vigilance tests for providing ground truth while wearing commercially available glasses equipped with EOG sensors. We found a statistically significant, positive correlation between BF and RTs meaning that blink frequencies increase along with reaction times (*i.e.*, slower reflexes) over the day.

The main contribution of this work is as follows:

- We present results of a 2-week in-the-wild study showing the connection between increasing reaction times due to the homeostatic process and rising blink frequencies throughout the day.
- We present a model which allows continuously recorded EOG data and the resulting eye blink frequencies to predict fatigue level changes in everyday settings.
- We publicly release our dataset along with Psychomotor Vigilance Task (PVT) measurements and self-assessments, which amount to 2,860 hours of EOG data and 1,047 ground truth assessments from 16 participants collected in-the-wild [1].

2 BACKGROUND

The work presented here builds upon research from the fields of cognitive psychology, wearable sensing, and cognition-aware systems.

Alertness and Fatigue

Alertness is a psychomotor quality, which describes our readiness to respond to stimuli but also plays a role in our higher cognitive functions affecting our productivity, decision making [50], and memory [63]. Throughout the day, our alertness levels succumb to systematic changes [29]–subject to *circadian rhythms*–resulting in phases of high alertness, during which we can perform tasks with high precision and phases of low alertness, during which we have a hard time concentrating [10, 59]. Awareness of these fluctuations enables us to gain an understanding of productive hours during the day, but also avoid critical work prone to accidents due to fatigue [22]. “Fatigue and subjective sleepiness [...] express the relationship that exists between alertness and performance during wakefulness on the one hand and sleep on the other hand.” [66]. Fatigue, therefore negatively affects alertness, resulting, for example, in slower reaction times.

Our biological clock regulates our cognitive and physical performance during waking hours creating circadian rhythmicity. Consequently, levels of cognitive performance, fatigue and alertness are changing throughout the day. Knowledge of these rhythms can provide so-called cognition-aware systems with a “window into [our] mind” [67] allowing them to adjust to the user’s current cognitive capacities. Alertness

[1] https://github.com/tagbenja/Fatigue_EOG_Raw

levels can be assessed through subjective as well as objective measures: subjective measures commonly refer to self-assessments, such as the Karolinska Sleepiness Scale (KSS) describing the perceived state of drowsiness [4] and the Stanford Sleepiness Scale (SSS), which inquires an imminent sleep onset [34]. Such assessments, however, do not only require prompting users throughout the day and therefore cause interruptions and can be prone to subjective biases [28, 66], but are often not accurate, because of impaired cognitive performance due to sleep deprivation [5]. Objective measures, on the other hand, can be based on different performance assessment tasks, such as search and find tasks and reaction time tests [66].

One of the most widely used tests for measuring alertness is the PVT, which measures the reaction times of users to visual stimuli appearing at random time intervals [23]. The PVT has since been adapted to mobile phones [43] and integrated into cognitive assessment toolkits [26]. In our work, we use the PVT to establish a ground truth for alertness levels throughout the day.

Sensing In-The-Wild

Traditionally, investigations into people's alertness levels required constrained lab studies, cumbersome examinations, such as rectal temperature monitoring, or powerful medical devices, such as Functional Magnetic Resonance Imaging (fMRI)s [35, 45, 48]. With new developments in mobile sensing, medical grade equipment is not crucial anymore for the research of cognitive states as Pielot *et al.* [51] have shown in their work of identifying states of boredom through mobile phone usage. The resulting models were subsequently deployed in applications to trigger reading as well as learning reminders throughout the day [27]. In order to put findings to a test and validate lab-results in unconstrained settings, in-the-wild (or in-situ) studies have been shown to be the right testbeds [37, 53]. Nevertheless, due to the strong fluctuations, individual characteristics, and masking factors influencing measures of cognitive performance, noisy data recordings are a common issue [13, 54]. For example, an increased intake of caffeinated drinks can mask a person's sleepiness at a certain point in time. In-the-wild studies do not allow to control for influencing factors in the natural environment and, therefore, require big groups, long-term experiments, or mathematical models to control for sources of noise.

Abdullah et al. [3] recently introduced a machine model to predict alertness levels based on phone usage. The caveat of such models is that they require users to actively use their phones. Unobtrusive, but permanent sensing would provide a more holistic picture of people's alertness levels throughout the day and prevent unnatural behavior.

Eye Data and Electrooculography

Eye movement, as well as Electroencephalography (EEG) data, has been shown to provide insights into cognitive states and processes [15, 47]. EOG sensors, in particular, use the electrical activity caused by eye movements, which we utilize in our work to record eye blinks for monitoring alertness levels. Even though the number is highly flexible and depends on subjective conditions and environmental factors, eye blinks occur at an average of 15-20 times per minute [9]. Eye-blinking is not only responsible for lubricating and cleaning the eyeball but is also directly related to neural activities [31, 59]. Studies have demonstrated that about one-third of our natural eye blinks are sufficient to fulfill the cleaning function [41], meaning that the remaining blinks serve a different purpose [60]. Recent studies have shown that when humans are engaged in social communication situations their blink patterns change significantly. Especially states of arousal, cognitive engagement, and emotional changes have a direct impact on the blink frequency [49]. Moreover, blink rates have been shown to increase with raising fatigue levels while eye movement speed decreases and blink durations get longer [55]. Recent works, such as by Haq *et al.* [33] present highly accurate methods to detect eye blink rates of drivers indicating drowsiness and fatigue levels. While the application case is limited to the user being in front of the stationary camera setups, the necessary image processing, and computer vision algorithms [46] require considerate computational complexity [54]. Less cumbersome setups are made possible by mobile systems, which often rely on commercially available head-mounted infrared cameras installed in eye trackers [42] or on infrared reflectance sensors [19]. While the utilized bright infrared light bears an inherent risk of irritating the eye through the emitted heat if not properly set up, different works have also shown that these systems are likely to produce faulty measurements because of changing light conditions [40, 64], rendering their in-the-wild use as not feasible.

EOG-based systems have been successfully used for activity recognition in the past [14], but required sensors to be attached to the face and connected with a computing unit through cables, which rendered the setup rather intrusive. The possibility of integrating EOG sensors into regular prescription glasses makes this technique feasible to consistently track eye movements throughout the day in an unobtrusive way. EOG is immune to any form of light changes enabling eye movement measurements in well-lit (outside, daytime) as well as in dark environments (inside, nighttime). EOG utilizes the electrical potential difference between the cornea (+) and the retina (-), which changes when the eyes move. When closing the eyelid during a blink, eyes perform a characteristic nose- and downward oriented motion that can be measured by electrodes correctly placed around the eyes and nose [20]. EOG offers a robust, low-power sensing solution capable of monitoring complex consecutive eye movements, rendering it ideal for recordings in everyday situations as well as a using it as a possible input modality for HCI and ubiquitous computing applications [13, 62].

Cognition-aware Systems

With the technique presented in this paper to unobtrusively and frequently measure alertness levels throughout the day, we aim to advance the development of systems that become cognition-aware [13, 24, 28].

Such systems are capable of adjusting user interfaces to the user's current cognitive capacities, therefore avoiding frustration and boosting productivity levels [25].

The tracking method presented in this paper paves the way towards an unobtrusive, yet continuous assessment of alertness levels to inform systems about the user's cognitive state.

3 METHOD

The aim of this study was to continuously record EOG data throughout the day and, hence, to investigate the correlation between alertness fluctuations and blink frequencies. We, therefore, used an in-the-wild approach to record a data set that would enable us to identify fatigue changes in an unconstrained, everyday setting. We adhere to Van Dongen and Dinges' definition of *fatigue* as a "loss of desire or ability to continue performing" [66]. Rising fatigue levels, therefore, coincide with declines in alertness and cognitive performance [10], which can be measured by investigating changes in RT [55]. To validate the alertness level ground truth we created a mobile app, which periodically prompted participants to complete a sequence of PVT to collect reaction times as measures of alertness together with the time of day. In addition to this, the app also collected data on participants' sleep patterns, self-assessed sleepiness, and naps as well as caffeine intake (see Figure 2).

Apparatus

We adapted the mobile toolkit by Dingler *et al.* [26] to collect our ground-truth data. The toolkit based on Android features a task battery enabling the assessment of alertness and different higher cognitive functions. Since the PVT has been shown to provide the greatest amount of data points and most accurate alertness measures, we limited assessments to this one and left out the other two task types provided. For collecting the EOG data, we used off-the-shelf J!NS Meme[2] [38] glasses. These are sensing devices equipped with EOG sensors around the nose for measuring eye movements

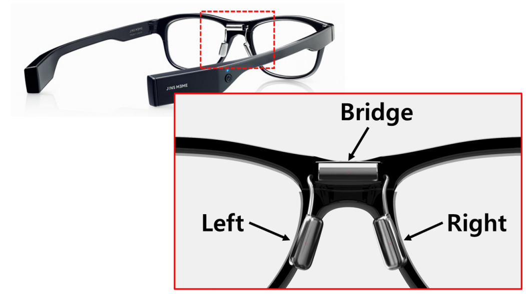

Figure 1: J!NS Meme glasses and closeup of the Electrooculography (EOG) electrodes.

and blinks, motion sensors (accelerator and gyroscope) for recording head movement and postures, and a Bluetooth LE (low energy) module for data transmission (Figure 1). We used the academic version of the glasses with a 50 Hz sampling rate and access to raw data for all EOG and Inertial Measurement Unit (IMU) recordings. Together with the J!NS Meme devices, we handed out Xiaomi mi4c smartphones to our participants, which recorded the EOG data as well as contained and triggered self-assessments every two hours ($\pm 20 minutes$). After each self-assessment, the PVT commenced with 10-15 rounds with random delays of 2-8 seconds between visual stimulus onsets resulting in test lengths of 20-120 seconds plus the individual RT for each round. The app recorded reaction times as well as the number of false attempts, such as taps that were made prematurely (faster than 100 ms) or too late (later than 3000 ms) in order to remove outliers and noise. A notification service running in the background made sure that participants were notified of the next pending survey in 100 - 140 minutes intervals. The goal of this implementation was to spread out measurements and collect a representative sample of fatigue measures throughout the day. If a notification was not immediately responded to, the application sent a new notification every five minutes until the participant finished the survey. To avoid sleep interruptions, we enabled a pausing function that stopped the notifications for a number of hours selected by the user. In cases where the user woke up earlier than planned, she could manually start the first/next survey. All recordings were stored locally on the phone and processed after each participant finished their study.

Participants

Through professional networks and university mailing lists, we recruited 16 participants (7 female) with a mean age of 28 years ($SD = 5.03$). All participants had normal or corrected to normal visual acuity, had no clinically significant problems, or were taking fatigue-inducing medication throughout the time of the study. The attendants who finished the full course of the study were compensated with JPY3,000 (*i.e.*, ca. 30 USD).

[2]https://jins-meme.com/en/

Procedure

We invited participants to our lab for an initial briefing session, where we introduced them to the purpose and procedure of the study and explained the functionality and controls of the smartphone and EOG glasses as well as how to charge either. All instructions were additionally handed out in written form before participants were asked to give their written consent. The study instructor then set up the equipment and demoed the different parts and functions of the smartphone application (Figure 2). Each briefing session took between 45 and 60 minutes. Necessary charging devices were provided as well.

The study ran for 14 days during which we asked participants to wear the J!NS Meme throughout the entire waking day, *i.e.,* from the moment they woke up until the time they went to bed, except for times of showering, bathing, swimming or other activities that would cause a risk to the user or the device. In the morning, the participants had to connect the glasses to the official J!NS Data Logger, an app installed on the phone, and disconnect it in the evening. Since the one battery charge of J!NS Meme is officially stated to last up to 16 hours, we advised all participants to charge the devices overnight in order to avoid possible recording interruptions due to running low on power. The EOG sensors integrated into the glasses' nose pads and bridge, permanently logged the eyes' EOG potentials throughout participants' regular daily activities. Data collection commenced the morning following each briefing. The app contained three different surveys (see Figure 2): upon the first launch of the app, participants were asked to provide demographic information. Every morning, the app triggered a first survey asking participants to indicate their wake-up time, the number of hours slept and had to evaluate the sleep quality on a scale from 1 (=poor) to 5 (=very good). Whenever participants clicked on a notification or opened the app, they were asked to fill in a brief momentary assessment about whether they had had a nap or any caffeinated drink within the time frame since the last survey. Further, they were asked to assess their current level of sleepiness on the KSS from 1 (=extremely alert) to 9 (=extremely sleepy) [39] (see Figure 2).

4 RESULTS

The goal of this study was to establish a relationship between eye blink frequency and fatigue changes throughout the day. Therefore, we need to verify the ground truth, *i.e.* changes in RT. Additionally, we look at different influencing factors, such as caffeine intake and sleep.

Over the course of the 14 days study, the 16 participants responded to an average of 4.09 (SD = 2.01) assessment tests per day, which accounted for an average of 65.44 (SD = 28.1) assessment test per person, with a minimum of 24

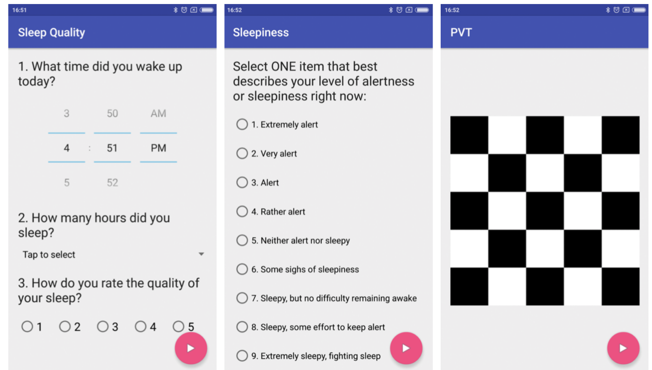

Figure 2: Android Application functions from left to right: daily sleep assessment, alertness self-assessment on the Karolinska Sleepiness Scale (KSS), and Psychomotor Vigilance Task (PVT) assessing reaction times

assessments and a maximum of 115 assessments, resulting in a total of 1,047 PVT assessments. Throughout their waking hours, participants were wearing the J!NS Meme glasses that permanently logged their EOG for a total of 2,860 hours of EOG raw data. Assuming an average of 16 wake hours per person and day, this would result in approximately 8.5 hours of EOG recordings per person and day. In order to be able to identify correlations between the reaction time and the blink frequency, we analyzed the 10-minute period of EOG data that directly preceded the respective assessment test. We chose this time window prior to the assessment test to avoid potential effects resulting from performing the PVT.

Ground Truth Validation: Performance Measures Throughout the Day

The homeostatic process dictates that the longer a person is awake, the stronger the sleep pressure becomes, resulting in a decrease of task performance. To validate our ground truth (*i.e.*, PVT measures) and detect systematic changes in the recorded performance, we fitted the data with a linear mixed model. Results obtained from the PVT were related to the deterioration of task performance throughout the day resulting in longer RTs as the day progresses.

Tukey outlier detection showed that both User 2 and User 10 fell outside the 3rd quartile in their average RT, with 794.07ms (SD = 223.29) for User 2 and 798.67 (SD = 204.59) for User 10, indicating either failure of equipment or failure in conducting the task. After removing the data from outliers and responses given during the night hours [1:00 a.m. - 6:00 a.m.], that accounted for 12 responses due to respective users' unusual wake hours, there were 937 observations left in our dataset for analysis. To be able to visualize the average trend of the development of RT values over time, we binned the average RT obtained through the PVT according to the hour

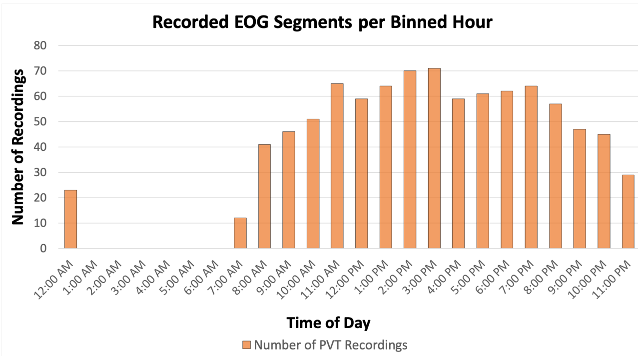

Figure 3: PVT measurements per binned hour with outliers (User 2 and User 10) removed.

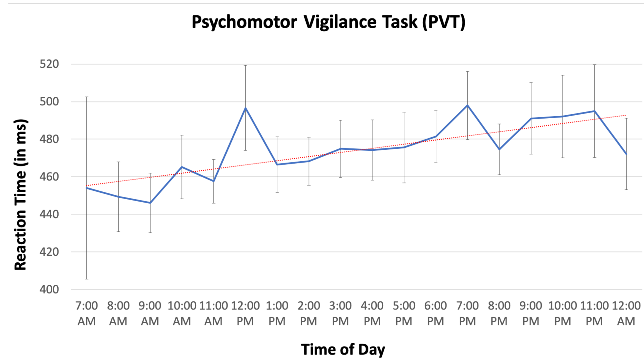

Figure 4: Visualization of variation in RTs across the day, as tested by a PVT (blue). The displayed trend line shows slowing of the reaction time over the day.

in which they fell across participants. This means that the bins have firm demarcations, which can lead to close times being put in different bins compared to further away bins. For instance, the pair [10:01 a.m. and 10:59 a.m.] would be in the same bin, whereas [10:59 a.m. and 11:00 a.m.] would fall in different bins. The binning also resulted in a greater error due to the different schedules of the users and lowered the effect of the homeostatic process depicted in our results. Nevertheless, rather than aiming at modeling the homeostatic process, we found it expressed in our results as a general increase in reaction times as the day progressed. The distribution of the number of responses can be seen in Figure 3. We furthermore fitted a trend line in Figure 4 in order to emphasize the slowing of the reaction time over the day. The binning has no effect on the results of the linear mixed model since this uses continuous data. Nevertheless, in order to visualize the collective trends and to enable the investigation of potential circadian rhythmicity, we chose hourly binning for our datasets.

We used recordings binned from 7:00 a.m. through 1:00 a.m. for our analysis. We fit a linear mixed model to the raw data with PVT average as the dependent variable and the fixed factors *time of measurement* (time of day in hours and minutes, with minutes, converted into the decimal system), *self-rated sleepiness*, *consumption of caffeinated drinks* in the preceding two hours, and *naps* in the preceding two hours. Caffeinated drinks and naps were treated as categorical variables and participating users were treated as a random factor. We corrected for multiple comparisons by using the Holm-Bonferroni procedure. Our analysis showed that time of assessment affected RT ($\chi^2(1) = 10.12, p = 0.002$), increasing it by about 2.34 milliseconds ±0.73 (standard error) per hour added throughout the day. The model was validated by a robust linear mixed model which accounts for the effects of outliers. The results were similar and the significant factor was retained. The increase in RT in our PVT recordings implies a deterioration of alertness with progressing time awake and constitutes the homeostatic process. An analysis for a significant influence of the circadian process on the RT, an omnibus test (ANOVA), did not reveal any significant results. Figure 4, however, shows patterns that coincide with findings of previous works [26], *e.g.*, the peak performance times in the hours between waking up and noon, the afternoon dip around 12 p.m. as well as the evening peak performance period in the 8 p.m. bin.

Test for a possible influence of caffeine and naps have not yielded any statistically significant effect on RT. Furthermore, self-assessments of sleepiness did not show significant relation to the depicted RT fluctuations. In summary, the analysis of the PVT recordings eventuated in a clear presentation of the impact of the homeostatic process on task performance throughout the day, rendering our ground-truth validated.

Eye Blink Frequency as an Indicator of Fatigue

In the following, we present our blink detection algorithm and its validation, and the blink data analysis and results.

Blink Detection

Before analyzing the collected EOG data for eye blink frequency changes, we validated our blink detection algorithm in a pre-study. We collected EOG data of one person in two different settings. In order to have a mobile setup that allows for in-the-wild recordings, we installed two Pupil Labs eye-tracker [42] cameras on the J!NS Meme frame, one arm on each temple so that the cameras can point at the wearer's eyes. This enabled us to record video of the eye movements synchronous with the EOG signals from the glasses' sensors without putting intrusive devices or sensors on the wearer's body. These videos were used to manually label eye blink events. The test person had to wear the modified spectacles once while sitting in our lab resting and once while taking a walk outside. We complied to Ding *et al.*'s [21] recommendation that eye blink recordings shall at least be five minutes

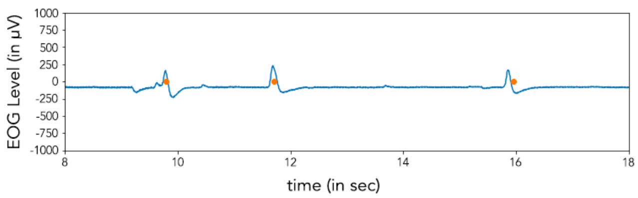

Figure 5: Plotted EOG data showing three detected (red dot) blinks in the vertical EOG.

long, because of eye blink frequency's natural tendency to fluctuate.

To estimate the eye-blink frequency, we apply a peak detection algorithm to the vertical components of the EOG collected with J!NS Meme. The transmission rate of the device is with 50 Hz half of the EOG sampling rate of 100 Hz, which results in a split of the original vertical value EOG_V into EOG_{V1} and EOG_{V2} packages. After combining the vertical EOG values EOG_{V1} and EOG_{V2} we used a low pass filter (Butterworth filter) to remove noise from the data. After filtering, the algorithm normalizes the data and moves a sliding window, with step size 0.01 sec over the data stream. Since eye blinks are characterized by a peak followed by a dip, the algorithm uses two thresholds *th_right* and *th_up_to_down* to detect blinks. *th_right* describes the height of the positive peak, i.e. the amplitude of the first and higher peak, and *th_up_to_down* describes the vertical distance between both peaks, i.e. the vertical distance between the highest point of the first peak and the lowest point of the second peak. In order to fully identify a blink, three conditions have to be fulfilled. Firstly, the height of the first peak has to be bigger than th_r. Secondly, the vertical distance (the difference between y-values of both peaks) between both peaks has to be larger than th_{ud}. Thirdly, the larger peak has to be followed by the lower peak, as can be seen in Figure 5. The detection algorithm was programmed in Python.

Validation

In order to validate the blink detection algorithm, and identify thresholds and window size for the blink detection, we compared the number of identified blinks in our resting and walking EOG data sets to the manually labeled blinks in the recorded video of the eyes. An eye blink in the recordings, shot by the eye tracking cameras, was counted when both eyes were simultaneously closed, with full closure not lasting longer than one second. We flagged in total 54 blink events throughout the five minutes of the resting state, and 84 blink events throughout the walking state. We achieved the most accurate blink detection with a th_r of 0.8, and a th_{ud} of 2.0. The best sliding window size was identified to be 0.34 seconds. Our algorithm detected 61 blinks/5min (12.2 blinks/min) in the EOG data of the resting person, and 81 blinks/5min (16.2 blinks/min) when the user was walking. This reveals a margin of error of +12.5% (+7 blinks) for the resting state, and -3.7% (3 blinks less) for the walking state.

Correlation Analysis

After running the blink detection algorithm with the validated threshold and window size settings, we fit a linear mixed model to the raw data with the RT obtained from the PVT readings as the dependent variable and BF. The participating users were treated as a random factor. We use the time codes of the PVT recordings to identify the times of assessments and extract the 10-minute EOG data segments that precede each assessment test. Pilot tests with different window sizes showed that the 10 min window yielded the best results as it is sufficiently large to collect a reasonable amount of data and small enough to still resemble a current, cognitive state. We removed 324 segments that did not contain any data, deriving from hardware issues, leaving 623 segments of valid EOG data. The 623 analyzed EOG segments yielded an average blink frequency of 11.4 blinks/min ($SD = 12.7$). Our analysis shows that BF changes with RT ($\chi^2(1) = 4.32, p = 0.001$), expressed in an increasing RT by about 1.64 milliseconds ±0.38 (standard error), and BF of +1 blink/min.

We corrected for multiple comparisons by using the Holm-Bonferroni method. The model was validated by a robust linear model which accounted for the effects of outliers. The results were similar, and the significant factor was retained. The results show that BF is an indicator for fatigue expressed in changes of RT (Figure 6, which coincides with the related literature [9, 59].

5 DISCUSSION

Results of our study show eye blink frequency to be a feasible indicator of an alertness-related performance decline throughout the day. Using unobtrusively collected EOG data, we validated the connection between an increase of reaction time (1.64 ms per hour) and a rising blink frequency (+1 blink on average per minute) as the day progressed.

While we collected alertness ground truth through assessment tasks, we used those to validate the connection between eye recordings and fatigue and showed the feasibility of unobtrusively monitoring fatigue levels. For future studies and applications that take into account users' alertness levels, the form factor of normal glasses provides the possibility to continuously collect fatigue data with users no longer being required to wear additional hardware and being released from following potentially interrupting self-assessment protocols.

The app we used for collecting our data allowed participants to delay and pause notifications for assessments, *e.g.*, for times of sleep, important work or school meetings in order to not excessively disturb their regular daily routines. As soon as these situations were finished, participants could manually respond to the assessments. To account for the resulting unbalanced design, we used linear mixed models that allow the statistical analysis of unbalanced datasets. The models detected statistically significant correlations between the time of day dependent performance measures (RTs), and BF obtained from EOG data sets being an indicator of fluctuations in performance. Compared to related work [66], we measured reaction times which were on average 100-200ms slower. This drift was proven systematic throughout all measurements based on a system lag of the smartphone/touchscreen used.

We validated the blink detection algorithm on only one user because we deemed the activity context (rest, motion) more important than individual signal differences. Eye blinks–as measures of a rather basic physiological nature–do not vary across humans to the same degree more complex behaviors and functions do (e.g., personality traits like extraversion). External validity can, therefore, be assumed to be comparably high as the resulting characteristic EOG pattern (a peak followed by a drop) is similar across users. Thus, we also chose a single sliding window size, based on the best detection accuracy in our validation study. The detected average BF of 11.4 blinks/min ($SD = 0.4$) is, nevertheless, lower than the reported average rates in healthy humans [9]. Whereas eye blink patterns are very susceptible to environmental factors such as humidity, lighting conditions, air streams, and the activity a person is currently engaged in, EOG is prone to noise from actions such as touches to the area around the sensors and motions such as jumps and walking. We attribute the lower measured BF to blinks not being detected by our blink detection algorithm when the data was very noisy. The blink detection algorithm showed these tendencies in the validation study with a -3.7% (3 blinks less) detection rate for the walking state.

Nevertheless, since the relative reaction times and BF changes are of importance to our analysis and the drift is systematic across subjects, we deem the differences compared to related works negligible. The patterns of the changes of RTs throughout the day are consistent with former studies on the homeostatic process. Even though tests for significant hourly differences in RT were inconclusive, our data indicates typical expressions of circadian rhythmicity, such as peak performance times in the morning and evening and the mid-day dip. We also tested for the correlations of caffeine, sleep, and self-rated sleepiness on the RTs, but our models did not detect any significant effects. We used the validated KSS for self-assessment of sleepiness, which was not identified as a significant predictor for objective performance measures. Since self-reports require a strong degree of self-reflection and introspection, their reliability is reportedly prone to conflicting findings [28, 66]. Related studies by Abdullah *et al.* [3] and Dingler *et al.* [26], also demonstrate only weak correlations between self-assessments and objective measurements.

The contribution of our work incorporates the design of an unobtrusive system, based on off-the-shelf devices capable of continuously collecting users' EOG data. From these, eye blink frequencies can be determined and considered as indicators of fatigue level variations caused by the latent homeostatic process. Furthermore, we verified the relationship between BF and RT in-the-wild over an extended period of time. This clearly differentiates our work from the large body of controlled studies relying on stationary setups, which were conducted over relatively short periods of time [7, 16, 68]. Compared to such controlled lab settings we had to deal with a comparably noisy dataset through the in-the-wild design. We collected data over 14 days while participants went about their daily lives, which allowed for a more representative data sample and compensation of extraordinary events, such as a night of bad sleep. This resulted in higher external validity but also more faulty data (29.2% of the 10min EOG windows were empty, due to, e.g., undiscovered disconnects). We averaged over multiple measurements for noise reduction, but for creating user-dependent models, future work will require larger datasets from individuals. Having validated the use of EOG in the form factor of normal glasses to track changes in fatigue levels across the day, our model now allows the development of a range of applications and research apparatuses for continuous data collection in-the-wild. The presented model allows for building context-aware systems that deploy blink rate measures as context variables to inform real-time systems, including monitoring fatigue levels to raise alarms in safety-critical situations or to proactively adjust settings (e.g., silencing notifications during fatigued periods). Additionally, we are releasing the data collected throughout this study as a public dataset in order to support further research and application development.

Application Scenarios

By integrating sensors in everyday devices, passive measurements of physiological signals are enabled, and the burden imposed on the user to self-assess is lifted. By using off-the-shelf EOG sensors, such as integrated into J!NS Meme, no additional hardware is needed to track people's alertness across the day. This regular and unobtrusive data collection enables a wide range of possible application scenarios. Context-aware systems can continuously support the physical and mental well being of their users by introducing recommendations and interventions. In times of onsetting exhaustion due to a demanding task, a reminder to take a break can help to replenish cognitive and physical resources. Systems that understand our biological clock and cognitive

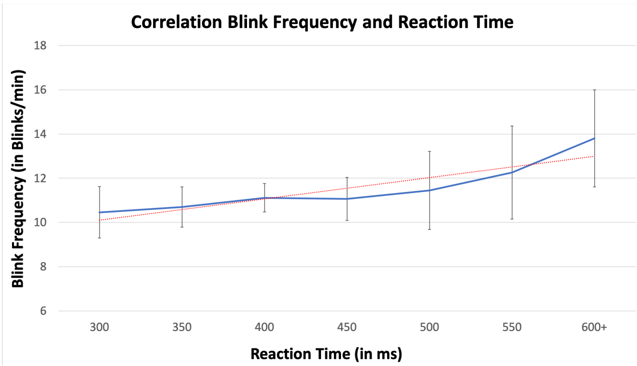

Figure 6: Visualization of correlation between blink frequency and reaction time (blue line). Linear trend is expressed in red line.

performance patterns can help us to schedule daily life activities, such as a hairdresser appointment in time slots that are usually defined by low alertness levels, whereas times of high alertness could be used to schedule an important work meeting. Momentary alertness assessments could inform user interfaces to adjust their complexity or task difficulty while long-term recordings of physiological data can be used to detect dominant patterns and build accurate models, for example for deciding on work shifts, and the best timing for vacation days. Additionally, group and teamwork could be timed in a way that members come together in periods with the highest average degree of alertness among all group members, allowing for flexible work times to become adaptive schedules responding to our biological rhythms. This system could also help to better adjust to new time zones in upcoming travels, and thereby help to tackle jet-lag more effectively.

A major advantage of such systems is the ability to detect in-situ changes in fatigue levels at any point in time. Cognition-aware systems with this ability can intervene when last-minute changes in fatigue patterns occur, *e.g.*, in a surgeon whose sleep was interrupted and who presents with unusually strong exhaustion. Especially, the recently promoted domain of semi-autonomous driving opens up a field of possible applications. Autopilots in cars and buses (also trains and plains) that are activated in as soon as the driver's fatigue level exceeds a certain threshold. This threshold can adjust to the current speed of the vehicle and environmental factors, by considering the impact of rain or snow on driving conditions. A different field where cognition-aware systems that react to our individual biological rhythms are promising to have a strong impact is the educational sector. We all are subject to our individual circadian rhythms [30]. Especially our established education system is still widely ignoring the fact that there are students with different chronotypes, who preferably and demonstrably perform better in the evening hours than in the morning hours. Fatigue-aware systems could help students study more efficiently and lower frustration for students and instructors, *e.g.*, by identifying increasing fatigue in a classroom among students. Teachers could be notified, and change the teaching method, give a break, or interact stronger with the students in order to increase attention. Especially the unobtrusive and passive data logging will ensure that distractions are widely avoided and no active engagement with anything but the learning material will be required from students.

Limitations and Future Work

One limitation of our setup is the susceptibility of the EOG signal to noise. Since the EOG sensors are attached to glasses' frames, their functionality is dependent on the proper placement of the glasses on the nose. The sensors are adjustable, and we made sure that every participants' pair of glasses was fitting well before the study started. Additionally, we ran different test runs with each participant in the preparation session to see how well the signal would be recorded. The J!NS Meme logger we used for our EOG recordings allows monitoring the recorded signal in real time. Nevertheless, when users touched their face, were moving quickly, were turning their head rapidly and even when they used their facial muscles intensely, the EOG signal became noisy, likewise observed by Rostamina *et al.* [54]. Removing the noise from the dataset and finding the right thresholds for detecting blinks properly was challenging. Additionally, even though we tested every glasses-phone connection several times before the study, random disconnects lead to the removal of ground truth data from our dataset.

Despite briefing all participants and asking them to try to avoid touching their face, wear too much make-up, and to check for disconnects, we were dependent on the diligence and compliance of our participants while at the same time not wanting to put too many constraints on them in order to preserve the character of the in-the-wild design. Limitations inherent to in-the-wild studies can also be seen as chances for generalizing and validating findings, as for example, the lack of control over the activities users were engaging in. A person who reads a lot will have a lower average BF compared to a person that is often engaged in conversations [9]. We believe that even though our data set was sufficient to control for these traits, we need a bigger set to find significance in the impact of coffee and sleep on fatigue levels.

Even though unobtrusive and passive sensing makes a step in the right direction, we still have to find ways to validate ground truth data without putting too much burden, especially on people who are skeptical of new technology or even afraid of using it. Especially when it comes to investigating CRs and their influence on fatigue and alertness, we have to find ways to include older adults in the group of participants, because age affects CR, too [6].

As mentioned earlier, blink duration is also directly related to fatigue [12, 58]. An analysis for changes in blink duration has not yet been included in our model. Whereas Bulling *et al.* [14] managed to accurately detect blink durations with a 128 Hz sampling rate, the available rate of our off-the-shelf device leaves the collection of data and subsequent analysis of blink duration changes for future work.

In the future, we are planning to combine the findings presented with other physiological sensing modalities, such as facial temperature changes for identifying cognitive load [1, 61], to identify more complex phenomena. For example, if we are able to detect fatigue together with high cognitive load, we can conclude that the person is dealing with complicated contents or tasks, whereas high fatigue and low cognitive load could be identified as boredom. By integrating these functions in everyday devices, such as glasses, we make further steps towards unobtrusive cognition-aware systems that enable us to detect complex contexts in everyday situations.

6 CONCLUSION

Alertness levels decline as the day progresses as a result of the homeostatic process. Measuring how this decline affects fatigue levels can help systems to proactively alert users to utilize phases of high alertness productively and refrain from accident-prone activities during fatigued phases. Moreover, the potential to continuously monitor data allows systems to detect changes in routines, such as travels to different time zones and can help to better cope with negative impacts. In this paper, we present results from an in-the-wild study showing the feasibility of using eye blink frequencies to detect an alertness decline throughout the day. By unobtrusively collecting EOG data through sensors integrated into normal glasses' frames, we validated the connection between an increase of reaction time and rising blink frequencies as the day progressed. The model along with the public release of the dataset collected allows future studies and applications to assess users' alertness levels without the need for additional hardware or potentially interrupting self-assessment protocols paving the way to building continuous, unobtrusive sensing for cognition-aware systems.

ACKNOWLEDGMENTS

We would like to thank all participants of our study. We further would like to thank Shoya Ishimaru from the University of Kaiserslautern and DFKI Kaiserslautern for his support and advice on eye blink detection. This research was supported by JST (Sakigake/Presto), Grant No: JP-MJPR16D4.

REFERENCES

[1] Yomna Abdelrahman, Eduardo Velloso, Albrecht Schmidt, Y Abdelrahman, and A Schmidt. 2017. Cognitive Heat: Exploring the Usage of Thermal Imaging to Unobtrusively Estimate Cognitive Load. *Proc.ACMInteract.Mob.WearableUbiquitous Technol* 13, 3 (2017), 33–33. https://doi.org/10.1145/3130898

[2] Saeed Abdullah, Elizabeth L Murnane, Mark Matthews, and Tanzeem Choudhury. 2017. Circadian Computing: Sensing, Modeling, and Maintaining Biological Rhythms. In *Mobile Health*. 35–58. https://doi.org/10.1007/978-3-319-51394-2_3

[3] Saeed Abdullah, Elizabeth L. Murnane, Mark Matthews, Matthew Kay, Julie A. Kientz, Geri Gay, and Tanzeem Choudhury. 2016. Cognitive rhythms: unobtrusive and continuous sensing of alertness using a mobile phone. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16* (2016), 178–189. https://doi.org/10.1145/2971648.2971712

[4] Torbjörn Åkerstedt and Mats Gillberg. 1990. Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience* 52, 1-2 (1990), 29–37.

[5] Paula Alhola and Päivi Polo-Kantola. 2007. Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment* 3, 5 (2007), 553–567. https://doi.org/10.1016/j.smrv.2012.06.007

[6] E. A. Antoniadis, C. H. Ko, M. R. Ralph, and R. J. McDonald. 2000. Circadian rhythms, aging and memory. *Behavioural Brain Research* 111, 1-2 (2000), 25–37. https://doi.org/10.1016/S0166-4328(00)00145-5

[7] Giuseppe Barbato, Gianluca Ficca, Giovanni Muscettola, Mariateresa Fichele, Michele Beatrice, and Franco Rinaldi. 2010. Diurnal variation in spontaneous eye-blink rate. *Psychiatry Research* 93, 2 (1 2010), 145–151. https://doi.org/10.1016/S0165-1781(00)00108-6

[8] Laura K. Barger, Brian E. Cade, Najib T. Ayas, John W. Cronin, Bernard Rosner, Frank E. Speizer, and Charles A. Czeisler. 2005. Worker Fatigue. *New England Journal of Medicine* 352, 2 (1 2005), 125–134. https://doi.org/10.1056/NEJMoa041401

[9] Anna Rita Bentivoglio, Susan B Bressman, Emanuele Cassetta, Donatella Carretta, Pietro Tonali, and Alberto Albanese. 1997. Analysis of blink rate patterns in normal subjects. *Movement Disorders* 12, 6 (1997), 1028–1034. https://doi.org/10.1002/mds.870120629

[10] Katharina Blatter and Christian Cajochen. 2007. Circadian rhythms in cognitive performance: Methodological constraints, protocols, theoretical underpinnings. *Physiology and Behavior* 90, 2-3 (2007), 196–208. https://doi.org/10.1016/j.physbeh.2006.09.009

[11] Alexander A. Borbély, Serge Daan, Anna Wirz-Justice, and Tom Deboer. 2016. The two-process model of sleep regulation: A reappraisal. *Journal of Sleep Research* 25, 2 (2016), 131–143. https://doi.org/10.1111/jsr.12371

[12] Gianluca Borghini, Laura Astolfi, Giovanni Vecchiato, Donatella Mattia, and Fabio Babiloni. 2014. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience and Biobehavioral Reviews* 44 (2014), 58–75. https://doi.org/10.1016/j.neubiorev.2012.10.003

[13] Andreas Bulling, Daniel Roggen, and Gerhard Tröster. 2008. It's in Your Eyes: Towards Context-Awareness and Mobile HCI Using Wearable EOG Goggles. *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)* (2008), 84–93. https://doi.org/10.1145/1409635.1409647

[14] Andreas Bulling, Jamie A. Ward, Hans Gellersen, and Gerhard Tröster. 2011. Eye movement analysis for activity recognition using electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 4 (2011), 741–753. https://doi.org/10.1109/TPAMI.2010.86

[15] Andreas Bulling and Thorsten O. Zander. 2014. Cognition-aware computing. *IEEE Pervasive Computing* 13, 3 (2014), 80–83. https://doi.org/10.1109/MPRV.2014.42

[16] Philipp P Caffier, Udo Erdmann, and Peter Ullsperger. 2003. Experimental evaluation of eye-blink parameters as a drowsiness measure. *European journal of applied physiology* 89, 3-4 (2003), 319–325. https://doi.org/10.1007/s00421-003-0807-5

[17] Julie Carrier and Timothy H Monk. 2000. Circadian Rhythms of Performance: New Trends. *Chronobiology International* 17, 6 (2000), 719–732. https://doi.org/10.1081/CBI-100102108

[18] Stefan Cohrs. 2008. Sleep disturbances in patients with schizophrenia. *CNS drugs* 22, 11 (2008), 939–962.

[19] Artem Dementyev and Christian Holz. 2017. DualBlink: A Wearable Device to Continuously Detect, Track, and Actuate Blinking For Alleviating Dry Eyes and Computer Vision Syndrome. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. Article* 1, 19 (2017), 1–19. https://doi.org/10.1145/3053330

[20] D Denney and C Denney. 1984. The eye blink electro-oculogram. *British Journal of Ophthalmology* 68, 4 (1984), 225–228. https://doi.org/10.1136/bjo.68.4.225

[21] Qiuping Ding, Kaiyu Tong, and Guang Li. 2005. Development of an EOG (Electro-Oculography) Based Human-Computer Interface. *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 7 (2005), 6829–31. https://doi.org/10.1109/IEMBS.2005.1616073

[22] David F Dinges. 1995. An overview of sleepiness and accidents. *Journal of sleep research* 4, s2 (1995), 4–14.

[23] David F Dinges and John W Powell. 1985. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers* 17, 6 (1985), 652–655. https://doi.org/10.3758/BF03200977

[24] Tilman Dingler. 2016. Cognition-aware Systems As Mobile Personal Assistants. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 1035–1040. https://doi.org/10.1145/2968219.2968565

[25] Tilman Dingler. 2016. Cognition-aware systems to support information intake and learning. (2016).

[26] Tilman Dingler, Albrecht Schmidt, and Tonja Machulla. 2017. Building Cognition-Aware Systems: A Mobile Tookit for extracting time-of-day fluctuations of cognitive performance. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (9 2017), 1–15. https://doi.org/10.1145/3132025

[27] Tilman Dingler, Dominik Weber, Martin Pielot, Jennifer Cooper, Chung-Cheng Chang, and Niels Henze. 2017. Language Learning On-the-go: Opportune Moments and Design of Mobile Microlearning Sessions. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, New York, NY, USA, Article 28, 12 pages. https://doi.org/10.1145/3098279.3098565

[28] Stephen H. Fairclough. 2009. Fundamentals of physiological computing. *Interacting with Computers* 21, 1-2 (2009), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

[29] Namni Goel, Hans PA Van Dongen, and David F Dinges. 2011. Circadian rhythms in sleepiness, alertness, and performance. In *Principles and Practice of Sleep Medicine (Fifth Edition)*. Elsevier, 445–455.

[30] David Goldstein, Constanze S. Hahn, Lynn Hasher, Ursula J. Wiprzycka, and Philip David Zelazo. 2007. Time of day, intellectual performance, and behavioral problems in Morning versus Evening type adolescents: Is there a synchrony effect? *Personality and Individual Differences* 42, 3 (2007), 431–440. https://doi.org/10.1016/j.paid.2006.07.008

[31] M. Haak, S. Bos, S. Panic, and L.J.M. Rothkrantz. 2009. Detecting stress using eye blinks during game playing. *10th International Conference on Intelligent Games and Simulation, GAME-ON 2009* April (2009), 75–82.

[32] Kerstin Hänecke, Silke Tiedemann, Friedhelm Nachreiner, and Hiltraud Grzech-Šukalo. 1998. Accident risk as a function of hour at work and time of day as determined from accident data and exposure models for the German working population. *Scandinavian Journal of Work, Environment and Health* 24, SUPPL. 3 (1998), 43–48.

[33] Zeeshan Ali Haq and Ziaul Hasan. 2017. Eye-blink rate detection for fatigue determination. *India International Conference on Information Processing, IICIP 2016 - Proceedings* (2017). https://doi.org/10.1109/IICIP.2016.7975348

[34] E Hoddes, V Zarcone, H Smythe, R Phillips, and WC Dement. 1973. Quantification of sleepiness: a new approach. *Psychophysiology* 10, 4 (1973), 431–436.

[35] Wytske A Hofstra and Al W de Weerd. 2018. How to assess circadian rhythm in humans: A review of literature. *Epilepsy & Behavior* 13, 3 (1 2018), 438–444. https://doi.org/10.1016/j.yebeh.2008.06.002

[36] J. A. Horne and L. A. Reyner. 1995. Sleep related vehicle accidents. *Bmj* 310, 6979 (1995), 565. https://doi.org/10.1136/bmj.310.6979.565

[37] Stephen S Intille, Ling Bao, Emmanuel Munguia Tapia, and John Rondoni. 2004. Acquiring In Situ Training Data for Context-Aware Ubiquitous Computing Applications. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 6, 1 (2004), 1–8. https://doi.org/10.1145/985692.985693

[38] Shoya Ishimaru, Kai Kunze, Katsuma Tanaka, Yuji Uema, Koichi Kise, and Masahiko Inami. 2015. Smart Eyewear for Interaction and Activity Recognition. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 307–310. https://doi.org/10.1145/2702613.2725449

[39] Bruno Jammes, Hassan Sharabty, and Daniel Esteve. 2008. Automatic EOG analysis: A first step toward automatic drowsiness scoring during wake-sleep transitions. *Somnologie* 12, 3 (2008), 227–232. https://doi.org/10.1007/s11818-008-0351-y

[40] Amir-Homayoun Javadi, Zahra Hakimi, Morteza Barati, Vincent Walsh, and Lili Tcheang. 2015. SET: a pupil detection method using sinusoidal approximation. *Frontiers in Neuroengineering* 8, April (2015), 1–10. https://doi.org/10.3389/fneng.2015.00004

[41] Craig N Karson. 1983. Spontaneous Eye-Blink Rates and Dopaminergic Systems. *Brain* 106, 3 (1983), 643–653.

[42] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. (2014). https://doi.org/10.1145/2638728.2641695 arXiv:1405.0006

[43] Matthew Kay, Kyle Rector, Sunny Consolvo, Ben Greenstein, Jacob O Wobbrock, Nathaniel F Watson, and Julie A Kientz. 2013. PVT-touch: adapting a reaction time test for touchscreen devices. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*. IEEE, 248–251.

[44] Simone M. Keller, Phyllis Berryman, and Eileen Lukes. 2009. Effects of Extended Work Shifts and Shift Work on Patient Safety, Productivity, and Employee Health. *AAOHN Journal* 57, 12 (2009), 497–502. https://doi.org/10.3928/08910162-20091116-01

[45] Nathaniel Kleitman. 1923. Studies on the Physiology of Sleep I. The effects of prolonged sleeplessness on man. *American Journal of Physiology* 66, 131 (1923), 67–92. https://www.physiology.org/doi/pdf/10.1152/ajplegacy.1923.66.1.67

[46] Alvaro Marcos-Ramiro, Daniel Pizarro-Perez, Marta Marron-Romera, and Daniel Gatica-Perez. 2014. Automatic Blinking Detection towards Stress Discovery. *International Conference on Multimodal Interaction* (2014), 307–310. https://doi.org/10.1145/2663204.2663239

[47] Milos Matousek and Ingemar Petersén. 1983. A method for assessing alertness fluctuations from EEG spectra. *Electroencephalography and clinical Neurophysiology* 55, 1 (1983), 108–113.

[48] Tamami Nakano. 2015. Blink-related dynamic switching between internal and external orienting networks while viewing videos. *Neuroscience Research* 96 (2015), 54–58. https://doi.org/10.1016/j.neures.2015.02.010
[49] Tamami Nakano and Shigeru Kitazawa. 2010. Eyeblink entrainment at breakpoints of speech. *Experimental Brain Research* 205, 4 (2010), 577–581. https://doi.org/10.1007/s00221-010-2387-z
[50] Conor O'Neill and Kriengsak Panuwatwanich. 2013. The impact of fatigue on labour productivity: case study of dam construction project in Queensland. *Proceedings from EPPM* (2013).
[51] Martin Pielot, Tilman Dingler, Jose San Pedro, and Nuria Oliver. 2015. When attention is not scarce - detecting boredom from mobile phone usage. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15* (2015), 825–836. https://doi.org/10.1145/2750858.2804252
[52] Alan T Pope, Edward H Bogart, and Debbie S Bartolome. 1995. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology* 40, 1 (1995), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3
[53] Yvonne Rogers, Kay Connelly, Lenore Tedesco, William Hazlewood, Andrew Kurtz, Robert E Hall, Josh Hursey, and Tammy Toscos. 2007. Why It's Worth the Hassle - The Value of In-Situ Studies When Designing Ubicomp. *UbiComp'07* 4717 (2007), 336–353. https://doi.org/10.1007/978-3-540-74853-3
[54] Soha Rostaminia, Addison Mayberry, Deepak Ganesan, Benjamin Marlin, and Jeremy Gummeson. 2017. iLid: Low-power Sensing of Fatigue and Drowsiness Measures on a Computational Eyeglass. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 2 (2017), 23:1–23:26. https://doi.org/10.1145/3090088
[55] Robert Schleicher, Niels Galley, Susanne Briest, and Lars Arnim Galley. 2008. Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired? *Ergonomics* 51, 7 (2008), 982–1010. https://doi.org/10.1080/00140130701817062
[56] Christina Schmidt, Fabienne Collette, Christian Cajochen, and Philippe Peigneux. 2007. A time to think: Circadian rhythms in human cognition. *Cognitive Neuropsychology* 24, 7 (oct 2007), 755–789. https://doi.org/10.1080/02643290701754158
[57] Christina Schmidt, Fabienne Collette, Yves Leclercq, Virginie Sterpenich, Gilles Vandewalle, Pierre Berthomier, Christian Berthomier, Christophe Phillips, Gilberte Tinguely, Annabelle Darsaud, Steffen Gais, Manuel Schabus, Martin Desseilles, Thien Thanh Dang-Vu, Eric Salmon, Evelyne Balteau, Christian Degueldre, Andre Luxen, Pierre Maquet, Christian Cajochen, and Philippe Peigneux. 2009. Homeostatic sleep pressure and responses to sustained attention in the suprachiasmatic area. *Science* 324, 5926 (2009), 516–519. https://doi.org/10.1126/science.1167337
[58] John A Stern, Donna Boyer, and David Schroeder. 1994. Blink Rate: A Possible Measure of Fatigue. *Human Factors* 36, 2 (1994), 285–297. https://doi.org/10.1177/001872089403600209
[59] John A Stern, D. Boyer, D.J. Schroeder, and and United States. and Civil Aeromedical Institute. 1994. Blink Rate As Measure of Fatigue: A Review. *U.S. Dept. of Transportation, Federal Aviation Administration, Office of Aviation Medicine* (1994), i, 12 p.
[60] John A Stern, Larry C Walrath, and Robert Goldstein. 1984. The Endogenous Eyeblink. *Psychophysiology* 21, 1 (1984), 22–33. https://doi.org/10.1111/j.1469-8986.1984.tb02312.x
[61] Benjamin Tag, Ryan Mannschreck, Kazunori Sugiura, George Chernyshov, Naohisa Ohta, and Kai Kunze. 2017. Facial Thermography for Attention Tracking on Smart Eyewear: An Initial Study. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2959–2966. https://doi.org/10.1145/3027063.3053243
[62] Benjamin Tag, Junichi Shimizu, Chi Zhang, Naohisa Ohta, Kai Kunze, and Kazunori Sugiura. 2016. Eye Blink As an Input Modality for a Responsive Adaptable Video System. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 205–208. https://doi.org/10.1145/2968219.2971449
[63] Maria Thomas, Helen Sing, Gregory Belenky, Henry Holcomb, Helen Mayberg, Robert Dannals, JR Wagner, David Thorne, Kathryn Popp, Laura Rowland, et al. 2000. Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity. *Journal of sleep research* 9, 4 (2000), 335–352. https://doi.org/10.1046/j.1365-2869.2000.00225.x
[64] Marc Tonsen, Xucong Zhang, Yusuke Sugano, and Andreas Bulling. 2015. Labeled pupils in the wild: A dataset for studying pupil detection in unconstrained environments. (2015). https://doi.org/10.1145/2857491.2857520 arXiv:1511.05768
[65] Tiinamaija Tuomi, Cecilia LF Nagorny, Pratibha Singh, Hedvig Bennet, Qian Yu, Ida Alenkvist, Bo Isomaa, Bjarne Östman, Johan Söderström, Anu-Katriina Pesonen, et al. 2016. Increased melatonin signaling is a risk factor for type 2 diabetes. *Cell metabolism* 23, 6 (2016), 1067–1077. https://doi.org/doi:10.1016/j.cmet.2016.04.009
[66] Hans P. A. Van Dongen and David F. Dinges. 2000. Circadian Rhythms in Fatigue, Alertness and Performance. *Principles and Practice of Sleep Medicine* 215 (2000), 391–399. https://doi.org/10.1002/ppul.1065
[67] Boris M. Velichkovsky and John Paulin Hansen. 1996. New Technological Windows into Mind: There is More in Eyes and Brains for Human-computer Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '96)*. ACM, New York, NY, USA, 496–503. https://doi.org/10.1145/238386.238619
[68] Mei Wang, Lin Guo, and Wen Yuan Chen. 2017. Blink detection using Adaboost and contour circle for fatigue recognition. *Computers and Electrical Engineering* 58 (2017), 502–512. https://doi.org/10.1016/j.compeleceng.2016.09.008