This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Theme Article: Personalized Pervasive Health

Inferring Circadian Rhythms of Cognitive Performance in Everyday Life

Benjamin Tag
The University of Melbourne

Tilman Dingler
The University of Melbourne

Andrew W. Vargo
Osaka Prefecture University

Vassilis Kostakos
The University of Melbourne

***Abstract*—Physical, mental, and behavioral processes of most living beings underlie cyclic changes, mainly governed by the day-night cycle. Investigations of these circadian rhythms have traditionally required constrained settings and invasive methods, such as repetitive blood testing and testing in sleep laboratories. Recent developments in pervasive technology, e.g., the proliferation of smartphones in our everyday lives, allow us to develop less intrusive ways to infer circadian rhythmicity in everyday settings. In this article, we present an overview of the current state of research, describe a mobile toolkit for collecting ground truth data on cognitive state fluctuations, and detail the implementation of a wearable system to unobtrusively detect alertness changes in the wild. Understanding and monitoring circadian rhythms will lead to the development of interventions to support mental health, physical health, and will ease the negative consequences of time shifts inflicted by jet lag or shift-work.**

How is it that at times the toughest challenge at work is engaging and joyful and we have no problems working concentrated for several hours, while at other times we have trouble focusing on a single paragraph of text?

It turns out that our alertness, attention, and vigilance are highly variable and follow systematic changes across the day.[1] They are subject to biological processes, which fluctuate in cycles

Digital Object Identifier 10.1109/MPRV.2020.2994914

of a single day (i.e., circadian: *circa*, about; *diem*, a day) or longer than one day (infradian). These so-called *circadian rhythms* are a result of adaptations to the diurnal alternation between day and night, a natural cycle. That was until electricity came along turning nights into days and allowing us to extend the boundaries of our inner rhythmicity on a grand scale.

In 1938, Nathaniel Kleitman of the University of Chicago studied these ancient, natural patterns by moving into a dark cave together with his student, Bruce Richardson. Deprived of natural light and any time cue, over one month in the cave, they attempted to force their bodies to conform to a non-24-hcycle.

By measuring their daily rhythm of core body temperature, Kleitman found an endogenously generated 24-h body temperature cycle, which seemed to ignore any attempted extension of his inner cycle.[2]

This cycle has evolved to synchronize times of activity of diurnal (daytime) and nocturnal (night-time) organisms ensuring active phases for food procurement and inactive ones for replenishing depleted energy. Consequently, a vast range of processes follow this rhythmicity, including our sleep/wake phases, moods and emotions, hunger and thirst, but also heart rate, blood pressure, body temperature, and changes in hormone production.

In practice, these rhythms are easily compromised by our tendency to override our predispositions through the consumption of caffeine, pharmacological interventions, or merely by sitting in front of a bright (blue light-emitting) screen. While cross timezone travel or unsuitable work shifts can send our inner clock into disarray, disruptions can also be of social nature: adhering to social demands by compromising biological rhythms, resulting in a so-called *social jet lag*.[3] If prolonged across an extended time period, the misalignment between our inner and factual clocks can have severe effects on our physical and mental well being. Repercussions can range from mild jet lag to sleep disorder and contribute to depression, coronary heart disease, and diabetes.[4]

Chronotypes, unique circadian profiles, place people on a spectrum between "early birds" (early risers) or "night owls" (late-night types). Diurnal variations in sleep propensity are generated by two underlying processes: the sleep/wake *homeostasis* and a *circadian process*.[5] The homeostatic process is characterized by an increase in sleep pressure while being awake, whereas sleep dampens the sleep pressure. The circadian process is influenced by external processes, such as the exposure to daylight.[6] While an increase of sleep pressure results in a decrease of cognitive performance during wake periods, the circadian process follows a sinusoidal pattern marking hours of the day when the sleep drive is particularly strong or rather weak. These two processes often either reinforce each other (at night) or cancel each other out (the reason why we are not fully alert right after waking up: the homeostatic process is at its peak while the circadian process is at its lowest).

Alertness is a crucial cognitive state, describing the degree of readiness to receive and process stimuli, while also impacting higher cognitive functions, such as decision making and memory.[7] Sleep deprivation leads to a deterioration of alertness and vigilance, a consequence of the developing *fatigue*. "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."[8] Thus, fatigue, i.a., negatively affects alertness, resulting, e.g., in slower reaction times (RT). Alertness levels fluctuate due to circadian and homeostatic influence, leading to periods of raised alertness where tasks are completed with higher precision and fewer mistakes, and phases of lower alertness resulting in decreased performance.

Kleitman's measures of circadian patterns relied on the body core temperature, obtained through invasive rectal temperature measurements. Similarly cumbersome, functional magnetic resonance imaging fMRIs have been used to monitor circadian rhythms. Recently, mobile sensors have become a feasible alternative to prior methods with the additional advantage of continuously providing measures throughout the day. The use of machine learning models has been proposed to utilize correlations between phone usage and circadian patterns.[9] Similarly, Tseng *et al.*[10] suggested a method to use the phone's front-facing camera to passively and unobtrusively monitor pupil information, which has been shown to be connected to alertness. These developments of sensors and algorithms open up new avenues for *circadian computing*, i.e., computing systems that take into account users' circadian patterns and provide functions to either manage these patterns or help detect and avoid disruptions.

We report on one study providing the blueprint for collecting ground truth through a toolkit, and a case study applying this method to build a robust model expressing the homeostatic process. We introduce an unobtrusive mobile solution that utilizes electrooculography (EOG). Both studies were conducted following an in-the-wild approach showing that today's consumer technology has the potential to infer everyday alertness fluctuations. Throughout this article, we will discuss results, limitations, opportunities, and implications of this approach.

Unobtrusive collection of sensor data will allow us to study circadian disruptions on a large scale and bring circadian computing closer to consumer applications. By collecting large data quantities on circadian patterns, we can start building individual models of alertness and fatigue, advance research on opportune times for the administration of drug treatments (optimizing delivery and efficacy), and to implement chronotherapy, i.e., applying circadian awareness to medical practice. Built into consumer applications, these models will help us schedule our day more effectively, but can also inform us when to undertake safety-critical tasks where high alertness levels are needed.

COGNITIVE TOOLKIT

We developed a toolkit to extract diurnal fluctuations of cognitive performance. It comprises short assessment tasks to be used in everyday settings and focuses on performance measures that quantify cognitive performance.

To cause as little disruption as possible to the daily routine of users, we limited each test's length to a maximum of 2 min. This conciseness of the tests came with the tradeoff of fewer collected data points, which can be accounted for by appropriate statistical methods.

Test Battery

To collect ground truth data on cognitive performances, we implemented a psychomotor vigilance task (PVT), a Go/No-Go task (GNG), as well as a multiple object tracking task (MOT) (see Figure 1). This enabled us to test higher cognitive functions, such as executive control and divided attention in addition to the RT collected through the PVT.

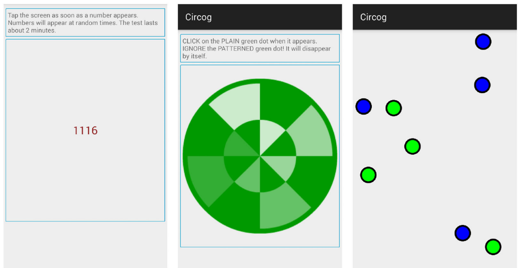

Figure 1. Toolkit comprises three tasks to measure alertness and cognitive performance variations across the day: A PVT (left), a go-no-go task (middle), and a MOT task (right).

Psychomotor Vigilance Task The PVT requires the user to react to a series of visual stimuli shown at random time intervals (2–6 s). It traditionally tests the subject's ability to sustain attention for 10 min,[11] but Roach *et al.*[12] showed that the RTs in a 90-s version of the PVT still presented medium to strong correlations to the findings from the longer version. To reduce intrusiveness in an in-the-wild scenario, we tested a very short version of the PVT (1 min) for its suitability to measure time-of-day dependent fluctuations.

Go/No-Go Test The GNG uses two clearly distinguishable stimuli to test the users' response inhibition capability. Each of the two stimuli is associated with one response requirement, i.e., either "go" or "no-go." Rather than only measuring RTs, the GNG also examines the ability to execute the cognitive function of deciding if a reaction is required.

Multiple Object Tracking Task The MOT is a sustained attention task that requires users to follow multiple moving objects. Divided attention is needed to correctly track and identify a subset of objects.

Study Procedure

To identify the efficacy of the implemented short versions of the PVT, the GNG, and the

MOT to detect diurnal changes in alertness levels in the wild, we conducted a validation study. All participants were asked to carry out each of the three tasks in their everyday life. Our dataset comprised of the *time of day* when the task was performed, *subjective sleep* and *alertness assessments*, as well as the *task performance* measures.

We developed (and publicly released https://github.com/Til-D/circog.git) an Android smartphone application that enables researchers to obtain raw measurement data from the phone's local storage. The application contains all three tasks and a notification scheduler as well as a logging service. For our initial validation study, we recruited 12 participants (4 female) with an average age of 24 years ($\mathrm{SD} = 2.67$). The study followed corresponding ethics procedures and all participants gave written consent.

Up to six times a day, the app triggered notifications to remind participants for performing all three tasks in a randomized order. Notifications were placed in the notification drawer until clicked or dismissed. Before everyday's first set of tasks, participants had to report their *wake-up time*, *hours of sleep*, and rate their *sleep quality* (1 = *very poor*, 5 = *very good*). Additionally, we asked participants to report if they had consumed any *caffeinated drinks* within the last hour and to assess their *perceived alertness levels* (1 = *super sleepy*, 5 = *super alert*). The study ran for two weeks, with a minimum required compliance rate of seven days.

Analysis

We analyzed our dataset investigating the efficacy of the test battery to identify cognitive performance fluctuations, the impact of sleep, caffeine, and accuracy of participants' assessments of their subjective alertness levels. The average participant performed tasks on nine days ($\mathrm{SD} = 3.9$). After removing all incomplete tasks, i.e., tasks that were not completed or for which self-assessments were missing, a total of 1098 completed tasks remained in our dataset (MOT: 367, GNG: 364, MOT: 367). Referring to the two-process model of cognitive performance fluctuations, we used two models to identify homeostatic and circadian changes.

Homeostatic Process The homeostatic process—a deterioration of cognitive performance with increasing time awake—implies that RTs should increase, i.e., the reaction to visual stimuli should slow down. Consequently, participants should become slower as the day progresses (and their sleep drive accumulates).

We fitted a linear mixed model to our data to determine the effect of time, *self-assessed alertness*, *caffeine intake* in previous hour, sleep duration, self-rated *sleep-quality* as well as the random *factor participant* on the performance measures *RT* (PVT and GNG), *false alarms* (GNG), and *misses* (MOT).

Circadian Process Circadian rhythmicity yields nonlinear performance fluctuations throughout the day. To minimize interruptions of participants' everyday life, we did not collect data at night, neither did we place participants on specific sleep schedules, which would have resulted in a disruption of the homeostatic process. Consequently, our findings do not fully resemble the sinusoidal shape of circadian rhythmicity of cognitive performance found in traditional laboratory studies.

Findings

For brevity, we only report significant results here. Sleep quality was overall assessed as positive on a 5-point Likert-style scale ($\mathrm{Mean} = 3.4$, $\mathrm{SD} = 0.8$), and sleep duration was reported within the range of regular sleep duration ($\mathrm{Mean} = 7.4\ \mathrm{h}, \mathrm{SD} = 1\ \mathrm{h}$). Consequently, variations in sleep quality and duration were too small to significantly impact cognitive performance.

On average, participants' RTs increased by 1.9 ms (± 0.7) ($\chi^2(1) = 6.7, p = 0.009$) per *hour of day* (see Figure 2). This indicates a statistically significant, slowing reaction to visual stimuli with increasing time awake echoing the impact of the homeostatic process on alertness. To identify circadian influence on the alertness levels, we analyzed the effect of the hour on performance. We found significant effects, especially between 1 P.M. and 2 P.M. (+54 ms SE) ($t(3173.6) = 3.3, p < 0.001$), with the mean RT being the highest between 2 P.M. and 3 P.M., indicating the characteristic, often called postlunch dip[6] in alertness. The reported RT values are the result of fitting our model to the collected data. Rather than measuring whether a task is

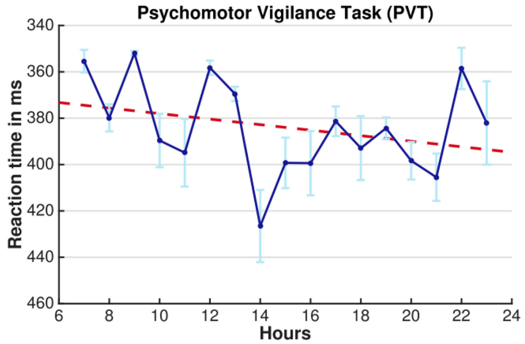

Figure 2. Performance variations across the day in blue: mean RTs from the PVT. The red line depicts a linear fit to the data. The ordinate axis has been inverted for the PVT to better communicate deterioration of performance. Error bars indicate the standard error of the mean. Model fitting was conducted on the raw data.

performed 10 ms slower or faster, we are using the measured variations as indicators for periods of low alertness. The data model yielded significant results during these periods, in which people show other cognitive changes coinciding with slower RTs. Instead, we fit our models on the raw data, i.e., RT values to single presentations of a stimulus rather than to summary statistics like means and standard errors.

We further found that caffeinated drinks, consumed within the hour before performing a task, had a significant impact on RTs, lowering them on average by 17.6 ms (± 13.7 ms SE) ($\chi^2(1) = 10.8, p = 0.001$). Participants were also able to determine times of high focus and alertness and identify them through self-assessments. The self-assessments were significant predictors for PVT RTs, i.e., with each point on the scale (i.e., higher alertness), participants' RTs decreased by an average of 10.4 ms ($\pm$4SE) ($\chi^2(1) = 5.8, p = 0.015$).

Implications and Application

This article shows that even data collected sparsely can be used to infer performance variations in-the-wild. While we limited each test's length to a maximum of 2 min, the task battery still required explicit user input to measure systematic performance fluctuations. The resulting statistical models showed that the PVT managed to elicit the homeostatic process as well as circadian variations of RTs.

A major tradeoff of the dispersed data collection is the difficulty to identify spontaneous fluctuations of cognitive performance measures. Utilizing users' alertness levels as an input modality could inform the design of scheduling applications, e.g., a calendar application, which predicts suitable time slots for tasks with varying cognitive demands. Data collected by this method would enable sufficient accuracy to predict alertness levels and allow the scheduling of activities throughout the day, e.g., meetings in the morning hours, exercise in the late-afternoon, and rather undemanding tasks in the post-lunch hour. Nevertheless, potential changes in lifestyle (e.g., sleeping patterns), seasonal changes (e.g., fewer hours of daylight in winter), and short-term circadian disruptions (e.g., jet lag) would require users to provide a new set of performance measures. To integrate these models into end-user applications, more implicit and continuous ways of gathering such data are needed.

BIOPHYSIOLOGICAL APPROACH UTILIZING EOG

Passive sensing solutions offer an alternative to collecting measures through active user input, such as alertness tests. Research by Stern *et al.*[13] showed that increased fatigue levels come with more frequent eye blinks informed our approach to use unobtrusive eye movement measures to detect fluctuations in cognitive performance.

EOG sensors measure eye movements and eye blinks. To infer alertness changes, we used off-the shelf glasses with integrated EOG sensors (J!NS Meme: https://jins-meme.com/en/) rather than attaching wet electrodes attached users' faces.

This approach enabled us to continuously log physiological signals while users can go about their everyday life. We adapted our mobile toolkit to contain only the PVT and self-assessment tools to collect ground truth data. Throughout a 14-day study, the participants of this second study were wearing the EOG glasses in their everyday life, allowing us to collect EOG data at 50-Hz sampling rate throughout their wake hours.

Eye Motion and EOG

EOG utilizes an electrical bi-pol around the eyes to measure eye motion. A high concentration

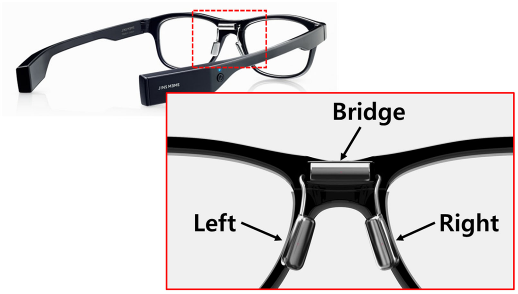

Figure 3. J!NS Meme glasses and closeup of the EOG electrodes.

of electrically active nerves in the retina synthesizes a negative electrical pole. As a consequence, the opposing cornea forms a positive pole, resulting in the so-called *cornea-retinal potential* CRP.[14] This electrical potential changes with every eye movement. Picking up these changes, EOG offers a low-energy, rather unobtrusive solution compared to approaches, such as fMRI and electroencephalography (EEG). We used off-the-shelf consumer glasses with a triangular dry-electrode setup without the need for any kind of adhesive (see Figure 3).

Apparatus

We relied on the PVT as the sole tool to collect RTs as alertness ground truth to limit interruptions to participants' daily routine to a minimum. Our smartphone application further issued brief surveys on *daily sleep behavior*, *caffeine-intake*, and self-assessment of *sleep-quality* and *alertness*. All data, assessments and EOG, were locally stored on the phone.

Procedure

The smartphone application triggered task reminders in the form of notifications every 2 h, ± 20 min for randomization. The purpose of this design was to distribute the ground truth data collection points as equally as possible over the whole day to obtain a representative dataset. The triggered notifications could not be turned off by the participants, but could be delayed for 5 min before notification was triggered again. We implemented a mechanism that allowed all participants to pause the notifications for a number of hours, e.g., sleep time.

We recruited a total of 16 participants (7 female) with a mean age of 28 years ($\mathrm{SD} = 5.03$). The article was approved by the University's ethics board, and all participants gave written consent. Participants needed to be between the ages of 20 to 40 years old and be capable of wearing the apparatus without endangerment of themselves or others. We employed this age restriction to control natural shifts in sleep behavior. Notably, we did not control for other lifestyle choices, such as employment or habits. This was deliberate to resemble conditions of a true in-the-wild study.

Results

We collected a total of 1047 PVT assessments. The average participant completed 65.44 ($\mathrm{SD} = 28.1$) assessments, approximating to 4.09 ($\mathrm{SD} = 2.01$) assessments per day. Through the EOG recordings, we obtained a total of 2860 h of raw EOG data. Our recordings averaged at 8.5 h of EOG recordings per person and day. To detect correlations between the EOG data and the assessment data, we focused our analysis on 10-min segments that directly preceded the time of an assessment test to avoid potential influence of the RT test on the blink frequency.

Ground Truth Performance Measures Tukey outlier detection identified two participants with significantly higher RTs of 794.07 ms ($\mathrm{SD} = 223.29$) and 798.67 ms ($\mathrm{SD} = 204.59$). RTs over 700 ms are more than twice as high as the average human RTs,[6] and indicate that the participant is impaired, not paying attention to the task, or that there is an equipment failure. The ratio of readings over 700 ms was 36 of 52 and 35 of 58, respectively. Visualizations of the recordings did not reveal any discernible pattern or point of equipment breakage. Consequently, we analyzed the ground truth PVT data from the remaining 14 participants. We then fitted the remaining observations with a linear mixed model, investigating the impact of *time of day*, *self-assessed alertness*, *caffeine intake*, and the additional factor *naps* on *RTs*. Our analysis identified a significant increase in RTs ($\chi^2(1) = 10.12$, $p = 0.002$), i.e., slowing of reaction speed to stimuli, of about 2.34 ms/h ($\pm 0.73\mathrm{SE}$) awake. The results show a clear impact of the homeostatic process on alertness, expressed in increasing RTs.

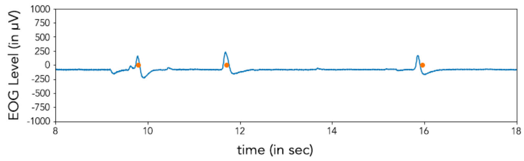

Figure 4. Plotted [EOG] data showing three detected (red dot) blinks in the vertical EOG.

To analyze our dataset for a reflection of circadian rhythmicity, we binned the recordings into 1-h segments. Even though we detected peak performance times in the morning hours, a post lunch-hour dip between 12 P.M. and 1 P.M., as well as an additional performance increase in the evening hours around 8 P.M., an analysis of variance did not yield any significant differences.

Our analysis furthermore did not show a statistically significant influence of *caffeine intake* or *naps* on the RTs. The same counts for self-assessments of *alertness* as a predictor for RT fluctuations. Different to what we reported for study 1, where we found a weak correlation between self-assessment and the performance measures, this article did not elicit any significant correlation. Often, self-assessments are prone to biases and distortions,[6] especially as increased fatigue can lead to such outcomes.

Blink Detection and EOG We used a peak-detection algorithm to detect the characteristic up-and-down movement the eye performs during a blink. This is expressed in a high peak followed by a dip in the vertical EOG (see Figure 4). After using a low-pass Butterworth filter to remove noise, our algorithm normalizes the data and moves a sliding window over the data identifying these characteristic patterns.

A major issue of eye blink patterns is the relative susceptibility to environmental influences, e.g., humidity levels, lighting, wind, but also the activity of a person. EOG signals are very sensitive to touches on the face and skin motions producing noisy readings. To account for these problems, we adjusted all sensors appropriately to each participant's facial anatomy and tested transmission rates for robustness. Nevertheless, in an everyday setting, the glasses we used could slide off of their placement on the face resulting in noisy or interrupted recordings. Face touching, rapid movements, hardware, and connectivity errors also led to data losses.

Findings

The requirement to interact with the PVT several times throughout the day renders our setup not fully unobtrusive and passive. This procedure was necessary to collect ground truth data to correlate changes in physiological signals to those in performance. Based on our findings in the previous study, we deployed the PVT to collect data that enabled us to gain a comprehensive depiction of the two-process model. We adjusted our experimental design so that collected samples were less scattered throughout the day, but could not elicit circadian rhythmicity in users' RTs. We could, however, correlate decrease in RTs showing the impact of the homeostatic process with the increasing average blink-frequency of our participants. Our analysis shows that blink frequency changes with RT $(\chi^2(1) = 4.32, p = 0.001)$. This is expressed in an increasing RT by about 1.64 ms $(\pm 0.38 \text{ SE})$, for each additional blink/min.

Our user-independent model is limited in versatility for real-life application cases, but nevertheless, shows that we can utilize physiological signals to infer cognitive state changes.

DISCUSSION

The presented toolkit and the proof-of-concept study using EOG data provide a pilot implementation for using unobtrusive sensing technologies to elicit models of systematic cognitive performance fluctuations in everyday life.

Our goal was 1) to show that cognitive variability across the day exists and 2) that cognitive variability can be tracked in less obtrusive ways, i.e., through eye blinks. Our studies present a proof-of-concept for implicitly collecting data for detecting alertness fluctuations. This approach builds the basis for future work, such as correlating similar measures to circadian variations, including user-device interactions, heart-rate variability, body temperature, and others.

A limitation of these studies is the relatively small number of participants. While this limits the overall generalizability, a contribution of this article is that it shows that in-the-wild studies using consumer technologies can detect known physiological patterns, such as homeostasis. And even though only a few restrictions were put on participant behavior, significant results could be produced. We asked participants to live their lives as they normally would, a key aspect of in-the-wild studies. We made sure that participants could withdraw from the study if they needed to start taking any prescription or allergy medicines and to record their caffeine intake.

Future studies at a larger scale would aim at improving the accuracy of our models. This would enable the detection of individual events, such as exceptional fluctuations (e.g., travel-induced jetlag), which can then be excluded from the data pipeline used to constantly refine a user-dependent model.

To put our results in perspective, study 1 showed an increase in RT by 1.9 ms/h of day PVT. The second study uses that same test at a similar level of control in the context of correlating eye blinks to fatigue. It reports 2.34 ms RT per waking hour added. Leaning on study 1—a validation study of three standardized tasks and their ability to detect changes in homeostasis and circadian rhythmicity—the EOG study extends this approach by correlating biophysical data that can be obtained continuously and unobtrusively. We expect that these two experiments and their respective successes will serve as a launchpad for much longer and larger experiments and implementations. We, therefore, released the toolkit for data collection and use in future studies similar to the one spearheaded by our EOG study. Such studies and replications will be important for verifying external validity and exploring the feasibility of applications based on circadian rhythmicity.

In studies like ours on cognitive performance fluctuations, potentially confounding factors need to be considered: 1) daily schedules and interruptions to these schedules, 2) factors acutely increasing and decreasing alertness, 3) environmental factors, and 4) natural acuity for psychomotor tasks. We controlled for these potentially amplifying or dampening confounds by requiring a minimum number of sample days. Additionally, we accounted for seasonality and geographic location of the experiments, which were kept consistent within each study. Personal factors are largely mitigated by the linear effect models used, which accommodate for variability between subjects. Factors, such as positive personal events, would add noise to the data, and if occurring regularly, would reduce the effect size that we observed, rather than amplify or cause it.

RESEARCH AGENDA

Our models describe general effects of a mix-gender, mix-age, mix-type form. To develop effective applications individual differences between users should be considered. We ran our studies during the summer months in Germany and Japan to guarantee similar natural light conditions. To design applications that support each individual user effectively, including different chronotypes, systems should collect data continuously and improve prediction and recognition accuracy of, e.g., opportune moments to work-out, study, or sleep.

Since it is barely feasible to attach EOG sensors to users' faces permanently, e.g., when swimming or working out, a multimodal approach could fill the gaps in the dataset: a combination of wearable sensors and smartphone interactions would potentially increase the data density and distribution. As heart-rate sensors in smartwatches are virtually standard these days, new devices include a variety of biophysical sensors with the potential to track systematic fluctuations. As studies have found that distal skin temperature decreases with increasing alertness,[15] wrist-recordings could be used to fill gaps of other data sources. Together with electrodermal activity, the measurement of skin conductivity, which correlates with human distress/calmness,[16] we can monitor a more comprehensive context of each user.

Smartphones can be used as the central computing unit, within such a sensor network. General user interaction opens up a vast range of relevant data: e.g., spelling errors and typing speed could be indicators for alertness fluctuations. Moreover, smartphones offer a variety of games that use different mechanics and game performance features, which can also be utilized as input modalities. Comprehensive and continuous data logging and analysis can run while users go about their everyday lives. The sum of these metrics might allow for prediction of highly individual cognitive performance models and would enable applications serving a wide range of purposes.

FUTURE APPLICATIONS

Peak cognitive performance is crucial in many safety-critical situations and occupations, such as flight monitoring or long-shift schedules in medical professions. Our approach and its resulting models allow applications to be developed for such use cases. By analyzing the usage context, a list of feasible sensing methods can be derived in a first step. A subsequent validation study using the toolkit presented would provide the data necessary to create models, which can then be deployed together with the sensor in-the-field.

For instance, for a work shift planner, an alertness scheduler could manage the rotation of security personnel at airports. A real-time monitoring system could then measure the personnel's *in situ* alertness levels to minimize the risk of missing important cues when alertness deteriorates. For end-users, circadian applications could help them to better manage jet lag, overcome sleep disorders, and schedule activities with general well being in mind. Reflecting on productive hours, circadian-aware applications open up new possibilities for application developers to consider fatigue levels as a context-aware feature and adjust the content and interface accordingly.

CONCLUSION

In this article, we provided a blueprint of ground truth collection using our open-source cognitive toolkit[17] and conducting an in-the-wild study with biophysiological sensors to track changes of alertness throughout the day.[18] While alertness is only one cognitive process subject to our inner clock, we propose a similar approach for future studies using off-the-shelf devices, such as the Ouraring (https://ouraring.com/), J!NS Meme or the Apple watch.

Current systems barely consider variations in sleep/wake cycles and the related diurnal patterns of cognitive performance. Awareness of our inner clock will help technologies support users with a wide range of pathologies, including sleep disorders, diabetes, depression, anxiety, bipolar disorder, and schizophrenia.[3] With 50–70 million people in the U.S. experiencing some form of sleep disorder, the Center for Disease Control warned that sleep disorders are reaching nearly epidemic levels.[19] Stabilizing sleep patterns and related aspects of people's circadian rhythms, therefore, offer an effective strategy to battle the effects of circadian disruptions and subsequently foster well being. Using mobile sensor data in combination with statistical models will allow circadian systems to determine and track individual chronotypes and diurnal patterns. While the mere awareness of their internal patterns can help individuals manage their processes and potential disruptions better, subsequent systems can also aim at triggering interventions with the goal of aligning social schedules and technological demands better with peoples' inner clock.[20]

ACKNOWLEDGMENTS

This work was supported in part by the Future and Emerging Technologies (FET) programme within the 7th Framework Programme for Research of the European Commission under FET Grant 612933 (RECALL), and by JST (Sakigake/Presto) under Grant JP-MJPR16D4.

REFERENCES

1. N. Goel, M. Basner, H. Rao, and D. F. Dinges, "Circadian rhythms, sleep deprivation, and human performance," *Prog. Mol. Biol. Trans. Sci.*, vol. 119, pp. 155–190, 2014.
2. N. Kleitman, *Sleep and Wakefulness.* Chicago, Illinois, USA: Univ. Chicago Press, 1993.
3. T. Roenneberg, K. V. Allebrandt, M. Merrow, and C. Vetter, "Social jetlag and obesity," *Curr. Biol.*, vol. 22, no. 10, pp. 939–943, 2012.
4. I. N. Karatsoreos, "Links between circadian rhythms and psychiatric disease," *Front. Behav. Neurosci.*, vol. 8, pp. 1–5, 2014.
5. A. A. Borbély, S. Daan, A. Wirz-Justice, and T. Deboer, "The two-process model of sleep regulation: A reappraisal," *J. Sleep Res.*, vol. 25, no. 2, pp. 131–143, 2016.

6. F. Wahl and O. Amft, "Data and expert models for sleep timing and chronotype estimation from smartphone context data and simulations," in *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–28, 2018.
7. M. Thomas *et al.*, "Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity," *J. Sleep Res.*, vol. 9, no. 4, pp. 335–352, 1999.
8. H. P. A. Van Dongen and D. F. Dinges, "Circadian rhythms in fatigue, alertness and performance," in *Princ. Pract. Sleep Med.*, no. 215, pp. 391–399, 2000.
9. S. Abdullah, M. Matthews, E. L. Murnane, and G. Gay, "Towards circadian computing: 'Early to bed and early to rise' makes some of us unhealthy and sleep deprived," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 673–684.
10. V. W.-S. Tseng, S. Abdullah, J. Costa, and T. Choudhury, "AlertnessScanner: What do your pupils tell about your alertness," in *Proc. 20th Int. Conf. Human-Comput. Interact. Mobile Devices Services*, 2018, pp. 41:1–41:11.
11. D. F. Dinges and J. W. Powell, "Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations," *Behav. Res. Methods, Instrum., Comput.*, vol. 17, no. 6, pp. 652–655, 1985.
12. G. D. Roach, D. Dawson, and N. Lamond, "Can a shorter psychomotor vigilance task be used as a reasonable substitute for the ten-minute psychomotor vigilance task?," *Chronobiol. Int.*, vol. 23, no. 6, pp. 1379–1387, 2006.
13. J. A. Stern, D. Boyer, and D. Schroeder, "Blink rate: A possible measure of fatigue," *Human Factors*, vol. 36, no. 2, pp. 285–297, 1994.
14. A. W. North, "Accuracy and precision of electrooculographic recording," *Invest. Ophthalmol. Vis. Sci.*, vol. 4, no. 3, pp. 343–348, 1965.
15. R. Lok, M. J. van Koningsveld, M. C. M. Gordijn, D. G. M. Beersma, and R. A. Hut, "Daytime melatonin and light independently affect human alertness and body temperature," *J. Pineal Res.*, vol. 67, no. 1, pp. 1–10, 2019.
16. R. Zangróniz, A. Martínez-Rodrigo, J. M. Pastor, M. T. López, and A. Fernández-Caballero, "Electrodermal activity sensor for classification of calm/distress condition," *Sensors (Switzerland)*, vol. 17, no. 10, pp. 1–14, 2017.
17. T. Dingler, A. Schmidt, and T. Machulla, "Building cognition-aware systems: A mobile toolkit for extracting time-of-day fluctuations of cognitive performance," in *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, Sep. 2017, vol. 1, no. 3, pp. 1–15.
18. B. Tag, A. W. Vargo, A. Gupta, G. Chernyshov, K. Kunze, and T. Dingler, "Continuous alertness assessments: Using EOG glasses to unobtrusively monitor fatigue levels in-the-wild," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2019, pp. 464:1–464:12.
19. S. T. Moturu, I. Khayal, N. Aharony, W. Pan, and A. Pentland, "Using social sensing to understand the links between sleep, mood, and sociability," in *Proc. IEEE Int. Conf. Privacy, Secur. Risk Trust/IEEE Int. Conf. Soc. Comput.*, 2011, pp. 208–214.
20. A. M. Schroeder and C. S. Colwell, "How to fix a broken clock," *Trends Pharmacol. Sci.*, vol. 34, no. 11, pp. 605–619, 2013.

Benjamin Tag is a postdoctoral researcher and associate lecturer with the School of Computing and Information Systems, University of Melbourne, Melbourne, Australia. His research interests include human cognition and cognition-aware systems. He focuses especially on inferring cognitive state changes from biophysical signals. Contact him at benjamin.tag@unimelb.edu.au.

Tilman Dingler is a lecturer with the School of Computing and Information Systems, University of Melbourne, Melbourne, Australia. He builds and investigates cognition-aware computing systems that support information processing and augment human cognition. Contact him at tilman.dingler@unimelb.edu.au.

Andrew W. Vargo is a project assistant professor with Osaka Prefecture University, Sakai, Japan. His research interests include the analysis and creation of collaborative systems for health and well-being. Contact him at awv@m.cs.osakafu-u.ac.jp.

Vassilis Kostakos is a professor with the School of Computing and Information Systems, University of Melbourne, Melbourne, Australia. His research encompasses ubiquitous computing, human–computer interaction, social computing, and the Internet of Things. Contact him at vassilis.kostakos@unimelb.edu.au.