AlertnessScanner: What Do Your Pupils Tell About Your Alertness

Vincent W.-S. Tseng
Cornell University
Ithaca, NY, USA
wt262@cornell.

Saeed Abdullah
Pennsylvania State University
University Park, PA, USA
saeed@psu.edu

Jean Costa
Cornell University
Ithaca, NY, USA
jmd487@cornell.edu

Tanzeem Choudhury
Cornell University
Ithaca, NY, USA
tanzeem.choudhury@cornell.edu

ABSTRACT

Alertness is a crucial component of our cognitive performance. Reduced alertness can negatively impact memory consolidation, productivity and safety. As a result, there has been an increasing focus on continuous assessment of alertness. The existing methods usually require users to wear sensors, fill out questionnaires, or perform response time tests periodically, in order to track their alertness. These methods may be obtrusvie to some users, and thus have limited capability. In this work, we propose *AlertnessScanner*, a computer-vision-based system that collects in-situ pupil information to model alertness in the wild. We conducted two in-the-wild studies to evaluate the effectiveness of our solution, and found that AlertnessScanner passively and unobtrusively assess alertness. We discuss the implications of our findings and present opportunities for mobile applications that measure and act upon changes in alertness.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous; See http://acm.org/about/class/1998/ for the full list of ACM classifiers. This section is required.

Author Keywords

Alertness; Fatigue; Mobile; Sensing; Pupil Segmentation

INTRODUCTION

Alertness is the state of readiness that enables us to respond to stimuli. The state of being alert plays a major role in our daily activities, affecting how we learn, solve problems, and remember facts and events [45]. However, there are several factors that can affect alertness, including circadian rhythms, sleep patterns, and stimulants such as coffee, which cause alertness to fluctuate during the day.

Given the importance of alertness in daily life, researchers have studied many interventions designed to improve alertness, such as napping [21] and exposure to blue light [47]. Researchers have also explored how to optimize activities based on alertness level, such as the 90-minute focus technique [15]. However, in order to build effective solutions, it is crucial to understand how alertness changes over time. For this purpose, a variety of methods have been developed and used.

Many techniques used to assess alertness involves using questionnaires [7, 24, 25]. Although questionnaires are widely used, there are conflicting findings regarding the reliability [22, 46]. In practice, individuals may not be be able to reliably gauge their own alertness level and report that in a questionnaire.

Given the limitations of questionnaires as instruments to measure alertness, researchers proposed other objective methods. One method that has become popular is the Psychomotor Vigilance Task (PVT) [18], which assesses alertness by measuring reaction time. In PVT, a visual stimulus is presented to the subject at random time intervals. The subject is asked to respond as soon as she sees the stimulus (e.g., by pressing a button). Although there are mobile implementations of PVT [19, 27], the methods require between 1 and 10 minutes to complete, rendering it inconvenient and impractical to use several times per day.

Another approach that can be used to assess alertness is the measurement of physiological changes. Previous research shows that physiological signals, such as brainwaves, eye movements, and blinks, can be used to reliably assess alertness level. This can be done using devices such as eye trackers, electroencephalograms (EEG) [26, 33], and electrooculography (EOG) [8]. However, most of the solutions used to measure alertness through physiological signals are either too burdensome or require constant calibration, rendering them impractical to use in-the-wild.

MobileHCI '18, September 3–6, 2018, Barcelona, Spain

DOI: https://doi.org/10.1145/3229434.3229456

Aiming to measure users' cognitive states in-the-wild, researchers have developed solutions for predicting states such as boredom [37] and alertness [5]. Although the methods proposed are effective, they require users to actively use their phones to extract the features for the models. Given the high variability in how individuals use their phones, building a predictive model might require a personalized model and labeled data, which imposes an additional burden on the end user.

Due to the limitations of existing alertness-measuring solutions, one question that arises is how to measure a user's alertness with minimal or no effort from the user. In this paper we present a solution: assessment of alertness by measuring pupil size from pictures taken by smartphones, either manually or automatically. We developed a mobile application, called *AlertnessScanner*, and conducted two studies to evaluate the effectiveness of our method. In both studies, we found that the method can be reliably used to infer alertness and requires minimal effort from the user.

In this paper, our contributions are threefold:

- We developed a system that captures in-situ pupil information for assessing level of alertness in the wild. We also implemented a system that uses computer vision to automatically infer alertness. We have made this computer vision based feature generation and learning code available for the community [1].
- We conducted two in-the-wild studies to evaluate the effectiveness of AlertnessScanner. In the first study, AlertnessScanner achieved a mean RMSE of 63.72 ms using single picture taken manually; in the second study, it achieved a mean RMSE of 43.28 ms with a burst of pictures taken automatically. The results from both studies suggest that the method can be reliably used to assess alertness.
- Finally, we discuss the implications of our system and findings, and present some opportunities for applications that measure alertness changes and act upon them.

RELATED WORK

Subjective Measures

There are several commonly used subjective measures used to evaluate alertness. For instance, the Karolinska Sleepiness Scale (KSS) is a nine-point Likert scale that describes the subject's state of drowsiness [7]. Another example is the Stanford Sleepiness Scale, which is a seven-point Likert scale which ranges from "feeling active, vital alert, or wide awake" (score = 1), to "no longer fighting sleep, sleep onset soon" (score = 7) [24]. The Epworth Sleepiness Scale (ESS) is a simple questionnaire based on retrospective reports of the likelihood of dozing off or falling asleep in a variety of different situations [25]. Unlike these discrete scales, Visual Analogue Scale (VAS) [9] asks participants to mark a position on a 10-cm horizontal line in between two endpoints to indicate their sleepiness level.

These self-report measures provide a simple and quick way to assess alertness. However, there are conflicting findings about whether self-report is a reliable measurement for alertness. Some studies have shown that self-reports of sleepiness and self-assessments of performance capability are unreliable [22, 46], and the onset of fatigue occurs before participant perceives sleepiness [36]. Other studies found that subjective assessments align with objective measurements [11, 20]. Aside from reliability, the subjective assessment can result in significant user burden. As a result, it may be very difficult to collect data continuously.

Objective Measures

Reaction Time Test

Apart from measuring subjective feeling, a number of studies have proposed objective measurements based on reaction time. The Psychomotor Vigilance Test (PVT) [18] is a widely used reaction time test. Its initial implementation required a specialized hardware device. The device shows a visual stimulus at random time intervals and the user has to press a button as soon as the visual stimulus appears. Longer mean or median response time, more lapses and increased number of false responses all indicate decreased alertness [38]. The length of the test ranges from 1 to 10 minutes [19, 28]. The original implementation required specialized hardware, but Kay et al. developed an Android-based implementation of PVT [27], which enables in-situ measurement using consumer devices.

These methods provide an objective assessment of alertness and can be generalized to different participants. The major downsides to these methods are that they are tedious and time-consuming, and they may even induce more fatigue in participant during the test. Consequently, the number of measurements that can be conducted in a day and the granularity of the information on how a participant's alertness changes over time might be limited.

Physiological Measures

Physiological signals have also been widely used to assess alertness objectively. For example, Electroencephalograph (EEG) can be used to assess electrical activity in the brain which correlates with alertness level [26, 33]. Similarly, electormyograph (EMG) has also been used for inferring alertness level. It records electrical activities of muscle, such as forearm muscle or facial muscle. EMG measurement is low to moderate for people who are alert, and virtually non-existent for people who are drowsy [8]. Eye activity also reveals a lot of information about alertness state. Electrooculogram (EOG) has been widely used to track eye features that indicate alertness states including slow eye movement, increased blink rate and prolonged blink duration[30, 42]. However, these kinds of physiological methods require special and costly hardware (e.g. electrodes placed on the skin) which might not be feasible to be deployed in the wild for a long period of time.

To make eye-tracking based methods more scalable, a number of recent studies have focused on developing fatigue detection

[1] The code for computer vision based framework described in this work will be available : `https://vtseng@bitbucket.org/vtseng/alertnessscanner.git`

systems using smartphone sensors. Most of these methods aim to track alertness in the context of driving. For example, He et al. [23] used smartphones to monitor the frequency of a driver's head nod, head rotation, and blinking during simulated driving tests. WakeApp [6] continuously tracks the percentage of eyelid closure (PERCLOS) over the pupil over time to infer alertness states. These methods can only detect pronounced changes in alertness, namely staying awake and falling asleep, but not subtle changes in alertness.

Pupil size is another physiological indicator of alertness that can be objectively measured. The size of the pupil is controlled by two complementary nervous systems — the sympathetic nervous system and the parasympathetic nervous system. When a person is alert, the sympathetic nervous system causes the pupil to dilate to facilitate information intake. On the other hand, if a person is drowsy, the parasympathetic nervous system will take over and cause the pupil to constrict [12, 29, 51, 54]. Pupil size has been shown to be a reliable indicator of fatigue [36] and sustained alertness [10]. However, using pupil size to assess alertness in the wild requires handling of a number of computational challenges including efficient image processing and controlling for environmental light. In this paper, we aim to address these computational challenges and investigate the feasibility of using pupil size as an indicator of alertness in the wild.

ALERTNESS SCANNER

We developed an Android application that measures pupillary response (the change in pupil size) using pictures captured by the front-facing camera of a smartphone. In this section, we describe how our application works.

One method used by researchers to measure pupillary response is calculating the pupil diameter in pixels. This method is usually used when eye images are taken by pupillometer or eye tracker, in which the distance between the camera and the eye is constant. However, this is not the case for pictures taken by smartphones, since users can face their phones at different distances. Depending on the distance between the eye and the smartphone camera, the pupil size might differ even when the alertness level of the user remains the same.

In order to address the aforementioned issue, we use pupil-to-iris ratio (PIR) as a measure for pupillary response. As the human eye attains its full size at the age of thirteen, the iris diameter can be regarded as a constant, and the size of iris diameter in an image can be used as a reference [34]. Therefore, PIR is consistent regardless of the distance and angle from which the image is taken.

We developed an unsupervised computer-vision-based algorithm to extract the PIR measure. The method consists of eye detection, iris segmentation and pupil segmentation. A summary of the method is presented in Figure 1.

The first step of our method is detecting participant's eyes. This is accomplished with OpenCV's Haar cascade classifier [2]. After detecting the eyes, we employ a method to detect the iris. One issue of detecting the iris is that light reflection on

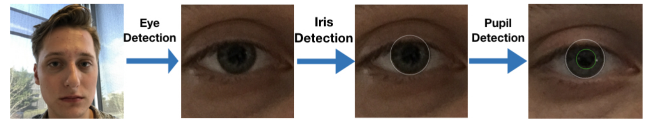

Figure 1. Framework for inferring alertness using pupil to iris ratio (PIR)

the eye results in a circular bright spot, which can be misclassified as the iris. To address this issue, we use a morphology transform: first inverting the color of the image and then performing closing [3] to remove specular reflections. After that, we employ Daugman's integro-differential operator to identify the location of the iris in the given image [17]. Daugman's integro-differential operator is given by

$$max_{r,x_0,y_0} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r} ds \right| \quad (1)$$

where G_σ is the Gaussian smoothing function with scale σ, and $*$ denotes convolution.

Given an input image, Daugman's integro-differential operator scans the entire image domain (x,y) to iteratively calculate the normalized contour integral of $I(x,y)$ along a circular arc ds of radius r and center coordinates (x_0,y_0). The circle that has the greatest contour integral derivative with parameter (x_0',y_0',r') is the geometry of the iris. In other words, (x_0',y_0') is the center and r' is the radius of iris.

After localizing the iris region, we can perform Daugman's integro-differential operator again to find the boundary of the pupil. Note that both pupil and iris are circular and concentric, and that the radius of the pupil is smaller than the radius of the iris. As a result, the pupil can be identified as the circular contour with the highest integral derivative within the iris region. In other words, when applied to iris region, the result of Eqn. 1 yields the center and the radius of the pupil. To prevent the occlusion of eyelids and makeup from affecting the iris and pupil localization, we only compute the contour integral in Eqn. 1 with θ in the range of $-45^\circ \sim 45^\circ$ and $135^\circ \sim 225^\circ$.

Finally, to calculate the PIR we use the two-box method [32, 35, 39, 41], in which two parallel rectangles are drawn along the same axis (horizontal axis in our case). We set one rectangle with width equal to the diameter of the iris, and one smaller rectangle with width equal to the diameter of the pupil . The ratio of the smaller rectangle's width to the bigger rectangle's width is the PIR. This method is shown to be reliable for estimating PIR when eye images are taken from various angles and at various distances from the camera.

FEASIBILITY EVALUATION

Before studying if it is possible to measure a user's alertness with AlertnessScanner, we decided to conduct a feasibility study to evaluate if the application could accurately detect the eyes, pupil, and iris of the user with front-facing camera of a smartphone. For this purpose, we recruited 2 participants,

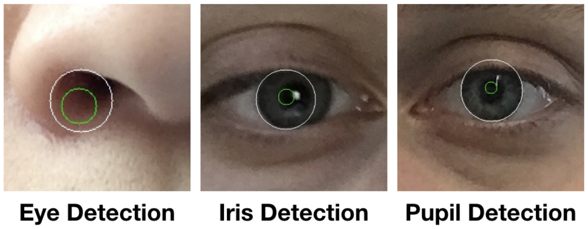

Figure 2. Examples of pictures in which the eye, pupil or iris were incorrectly detected

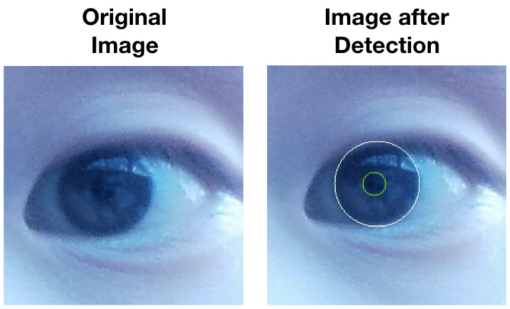

Figure 3. Example of picture in which the eye, pupil and iris were correctly detected

who collected data during one week using a Google Nexus 5 smartphone.

Both participants received several notifications per day at random time intervals, prompting them to take pictures of themselves using the frontal camera. A total of 50 pictures were collected by one participant, and 46 by the other. All pictures collected were then used as input for our algorithm. Our algorithm generated pictures with an annotation of the eye, pupil, and iris locations. After generating the pictures, one of the authors verified if the eye, pupil, and iris were correctly detected. This was done by manually checking the annotated images and comparing them to the pictures without the annotations. Figure 2 shows examples of pictures in which the eye, pupil, and iris were incorrectly detected, and Figure 3 shows an example in which everything was correctly detected.

The results of our feasibility evaluation were not very encouraging. The method was able to detect the eye, pupil, and iris correctly in only 36% of the pictures taken by one participant and 32% taken by the other. In order to improve the detection, we used one approach for facilitating the detection of the pupil and iris: removal of the infrared (IR) filter of the smartphone. The removal of the IR filter reduces specular reflections, making it easier to detect contours of pupil and iris even for people with dark eyes. Figure 4 shows an example of a picture collected with the IR filter and a picture without the IR filter. As the pictures show, the contours of the iris and pupil get more pronounced after the removal of the IR filter.

After removing the IR filters from two smartphones, the same participants that from the first data collection collected data over a week period using the modified smartphones. We noticed a large improvement in the detection. The method was able to correctly detect the eye, pupil, and iris in 71% of the pictures taken by one participant and 52% of those taken by

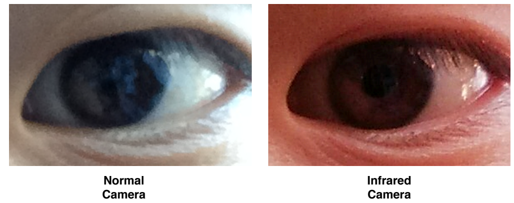

Figure 4. The image on the left was taken by normal phone camera, and the image on the right was taken by phone camera with the infrared (IR) filter removed. Removal of IR filter reduces specular reflections and makes it easier to detect pupillary boundary even for dark-eyed participants.

the other. By manually looking at the pictures in which the eye, pupil, or iris were incorrectly detected, we noticed some issues in the picture quality, including bad lighting conditions, blur, and even closed eyes. We realized that these issues could be addressed by allowing our mobile application to show a preview of the picture and by asking participants to check the quality of the picture before saving it in the app. After implementing these changes in the application and in the study protocol, we decided to conduct the first study, which is presented in the next section.

STUDY I

The aim of this study was to assess whether we could use pupil information extracted from phone-captured facial images to predict users' level of alertness. Based on the feasibility evaluation described earlier, smartphones with the IR-filters removed were used in this study to ensure the quality of images collected.

Participants

To recruit participants, we used snowball sampling and public noticeboards. During the initial meeting with a participant, we described the purpose and steps of the study in detail. The participants then signed a consent form. We gave each participant a Google Nexus 5 phone with AlertnessScanner, our customized data collection app. During the onboarding interview, we also demonstrated how the app works. The study lasted for three weeks. Participants were given compensation based on the duration they participated in the study ($5 for each week) and compliance rate ($0.5 for each completed task). All collected data were anonymized and the Institutional Review Board approved all the study procedures.

In total, 20 participants signed up for the three-week study. In this analysis, we filtered out data from five participants with compliance rate lower than 20%. The rest of the fifteen participants (5 blue-eyed, 1 green-eyed, 9 brown-eyed; 9 ranging in age from 18 to 22, 6 ranging in age from 23 to 27) had an average compliance rate of $63\% (S.D. = 22\%)$.

Data Collection

Our phone app collects data in two ways: (1) a sleep journal, which was collected once per day at 10 AM and (2) screen-unlock images, user-captured images, ambient light intensity, time of day, and Psychomotor Vigilance Task (PVT), which were collected every 3 hours between between 9 AM and midnight (12 AM). Participants were prompted to complete PVT by phone notifications. We will describe the data we collected in detail.

Sleep Journal

Sleep, sleep duration in particular, is another factor that affects a person's alertness during the day [50]. Therefore, we also collected information about daily sleep patterns. Participants were prompted to complete a sleep journal every day at 10AM. The journal entries included questions about when they went to bed and when they woke up, which later allowed us to calculate their sleep duration each day.

Screen-unlock Image

To help us assess the quality of images passively captured in a naturalistic setting, our app took a picture using the front-facing camera each time the phone was unlocked. As such, it did not require any user interaction. We did not use these images for developing models but instead used them to assess the feasibility of a completely automated alertness assessment system.

User-captured Image

In this study, we focused on assessing alertness using pupil information from facial images. To ensure good image quality and a sufficient quantity of images for modeling, participants were prompted to take an image of their face using the front-facing camera before the PVT session. Participants could see a preview of the facial image before the image was saved to ensure the quality of the image. The resolution of Google Nexus 5's front-facing camera is 1.3-megapixel (960 $\times$ 1280).

Ambient Light Intensity

Pupil size is affected not just by alertness but also by light. Pupils constrict when they are exposed to light and dilate in the dark. To account for this effect, we also recorded levels of environmental light while taking screen-unlock and user-captured images. For this, we used the light sensor in the phone.

Psychomotor Vigilance Task

We used Psychomotor Vigilance Task (PVT), an objective assessment of alertness, as our ground truth. PVT is a reaction-time task that measures the speed with which subjects respond to a visual stimulus [1, 28]. We employed PVT-Touch [27], which is a validated Android-based PVT implementation. During the task, participants are shown stimuli at random time intervals and asked to respond to these stimuli by touching the screen. These reaction times were recorded in milliseconds. Each PVT session was around 3 minutes long.

Data Processing

In this study, we extracted the following features: (1) pupil to iris ratio, (2) light intensity, (3) epoch, and (4) hours of sleep. Participants' pupil-to-iris ratios (PIRs) were computed based on the user-captured eye images. Given that time of the day influences a person's alertness level [14], we also used time information as a feature by categorizing the time when participant completed the PVT into one of the following epochs: midnight (12AM-6AM), morning (6AM-12PM), afternoon (12PM-6PM), and evening (6PM-12AM). It is worth noting that while we asked participants to self-report their hours of sleep, there are several existing algorithms that can automatically infer users' sleep duration [16, 48].

Data Analysis

Our goal is to model users' level of alertness using features described in the previous section. Median response time (MRT) is an indicator of alertness that has been widely used [38]. Using MRT enables us to measure and predict the level of alertness over a continuous spectrum. A higher value of MRT indicates lower alertness. As a result, we trained a regression model to predict users' median response time using support vector regression (SVR). To evaluate the models' performance, we used the following methods: 1) 10-fold cross validation, 2) leave-one-day-out cross validation, and 3) leave-one-person-out cross validation and compared the root-mean-square errors (RMSE) of the predicted MRTs. In addition, two different feature sets were used during each of the evalutions: the full feature set (PIR + light + epoch + sleep duration) and the reduced feature set (PIR + light). The reduced feature set was included in the evaluation so that we could investigate if our model would achieve comparable performance using features that can be collected automatically and without relying on phone-usage behavior to infer sleep. The intuition behind this is that information on sleep to some extent could already be revealed in users' pupil size [51].

Results

In total, 1,378 images were collected over the course of the study. The mean and standard deviation of participants' response time and the ambient light intensity are shown in Table 1.

Detection of eye, pupil and iris

We first focused on validating our algorithm for eye, pupil, and iris detection. As described in the section "Data Collection" the pictures were collected in two ways: i) manually, with the user pressing a button to take the picture and confirming the picture quality with an image preview; ii) automatically, by taking a picture whenever the user unlocked the phone.

First we looked into the precision of each module of the framework as shown in Figure 1. For this, we manually checked each image that passed through each step of the framework. The precision of each phase in our framework was very high — the precision of eye, iris, and pupil detection phases are 98.1%, 96.1% and 89.3%, respectively. This indicates that the implemented framework can effectively process these images. Furthermore, we also checked to what extent makeup adversely affects pupil segmentation. Among 240 images where participants had makeup, the algorithm achieved precision of 94.6%. This high precision might result from limiting the value of θ ($-45^\circ \sim 45^\circ$ and $135^\circ \sim 225^\circ$) as mentioned in the method section during the line integral.

We also analyzed the precision of our method using the pictures collected when the user unlocked the phone. This time, however, the results were significantly lower, with the precision of eye, iris, and pupil detection being 45.1%, 36.6%, and 10.7%, respectively. Since our method relies on correctly

Participant		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15
Response Time (milliseconds)	Mean	392.3	566.9	714.8	526.9	754.7	433.5	282.9	359.9	421.9	358.8	648.9	531.3	363.8	564.8	337.6
	S.D.	818.8	600.8	1085.1	794.0	990.6	638.0	71.3	175.5	419.4	173.4	841.6	374.9	174.7	435.3	136.4
Ambient Light (lux)	Mean	33.4	25.1	432.9	28.0	164.1	173.4	65.0	109.0	85.3	276.1	336.5	129.5	29.0	16.2	13.9
	S.D.	66.6	83.7	1607.3	47.2	733.3	1718.2	422.9	533.8	481.8	1989.7	1335.8	478.6	15.7	35.4	60.3

Table 1. Mean and standard deviation of PVT response time and ambient light intensity in StudyI.

identifying the pupil and iris to compute the PIR, we decided to continue the analysis using only the pictures collected manually.

Person-dependent Model

Our next step was to look at whether we could use an individual's own data to train a model that could predict their MRT. To this end, we trained a regression model on each individual's data using support vector regression (SVR), and evaluated the prediction performance using 10-fold cross validation. The mean root-mean-square error (RMSE) of the predicted MRT was 64.87 milliseconds (*Max*=138.12, *Min*=12.77, *S.D.*=36.65) and 72.65 milliseconds (*Max*=148.77, *Min*=12.06, *S.D.*=41.94) for models trained using the full feature set and the reduced feature set respectively. The results are comparable to the results from a previous study that employed phone usage as a predictor of response time [5]. This suggests that pupil information can be used to infer the level of alertness. More importantly, with pupil information and ambient light intensity alone, our proposed method are still able to differentiate between different levels of alertness.

In addition, we performed leave-one-day-out cross validation. Data from different days were iteratively used as test data, and the data from the remaining days were used as training data. The mean RMSE of leave-one-day-out cross validation was 63.72 ms (*Max*=149.16, *Min*=12.44, *S.D.*=39.86) and 66.56 ms (*Max*=199.11, *Min*=31.92, *S.D.*=42.12) for models trained using the full feature set and the reduced feature set respectively. The results are consistent with the results of the models trained on randomly shuffled individual data, which suggests that pupil size is a time-independent indicator of level of alertness.

Person-independent Model

Acquiring data from users in order to train personalized model might cause user burden and not be scalable. To this end, we trained generalized models using data from multiple users and evaluated the generalizability of the models. Given n users, during each iteration, a regression model was trained on data from $n-1$ users and tested on data from the remaining user. The mean RMSE is 105.39 ms (*Max*=268.72, *Min*=31.86, *71.80*) and 104.91 ms (*Max*=268.5, *Min*=31.92, *S.D.*=71.23) for models trained using the full feature set and the reduced feature set respectively.

Discussion

In this study, participants used a smartphone with a front-facing camera of 1.3MP without the infrared filter, and they were asked to take pictures of themselves manually using our app. After taking a picture, each participant had to do a PVT test to measure their response time. The results of the PVT test were used as a ground truth, and our goal was to predict the median response time obtained from the PVT tests using our pupil-to-iris ratio extraction method, along with other features collected passively.

The results of the study show that our models can be used to infer users' alertness using data collected from the camera and the light sensor in a smartphone. The mean RMSE of the predicted personalized model using 10-fold cross validation was 64.87 ms using 4 features and 72.65 using only 2 features. The mean RMSE of the generalized model was 105.39 ms using 4 features and 104.91 ms using 2 features. Both of the results are comparable to the results from a previous study that employed phone usage [5] as a predictor of user's response time.

Despite the encouraging results of this study, the protocol of the study still required effort from the user to collect the data. The data was collected when participants received notifications, and it required users to take pictures of themselves using our app. Furthermore, the removal of the IR filter facilitated the identification of the iris and pupil from the images, but it is not practical for users to remove the IR filters of their phones to use our solution. Because of these limitations, we decided to conduct a second study, in which the data was collected without the removal of the IR filter and without requiring any additional action by the participant. The second study is described in the following section.

STUDY II

In study I, we removed the IR filter and used a manual approach for picture collection in order to facilitate the detection of eye, pupil, and iris. However, these changes could be avoided if the smartphones had cameras with high enough resolution to make the contours of the eye, pupil, and iris very pronounced. Therefore, we decided to conduct a second study, in which participants used smartphones with a front-facing camera with much higher resolution (13 MP). Furthermore, in this study a burst of pictures were collected automatically when users unlocked their phones. By collecting several pictures instead of just one, we increased the chances of obtaining pictures with better quality.

Participants

In the second study, we followed the same recruiting procedure as we did in Study I. In total, 10 participants with an age range from 18 to 22 signed up for this study, and 2 of them dropped out due to not having enough time to continue the study. In the end, the rest of participants (4 blue-eyed, 4 brown-eyed) had an average compliance rate of 78%(*S.D.* = 34%).

Data Collection

The data collection scheme in Study II was the same as in Study I. We collected a sleep journal once per day and eye images, light reading, and PVT six times per day. The only difference was that in Study II, our app collected facial images in a completely automatic manner. Participants no longer needed to manually take pictures of their eyes. Instead, our app took 30 burst images of their eye (including the time when participant unlocked the screen to complete PVT). This increased the chance of capturing good-quality images.

Data Processing
The data processing also follows the same approaches we used in Study I, except for the computation of PIR values. In order to eliminate the artifacts caused by bad angles and poor lighting, we used the median PIR of all the images captured during each screen-unlock event.

Data Analysis
We employed the same approaches to modeling users' median response time. That is, using the full feature set (PIR + light + epoch + sleep duration) and the reduced feature set (PIR + light) to train person-dependent models and person-independent models.

Results
In total, 23,908 pictures were collected over the course of the study. Out of the 23,908 pictures, iris and pupil were correctly recognized in 17,206 pictures. The means and standard deviation of participants' response time and ambient light intensity are shown in Table 2.

Participant		P16	P17	P18	P19	P20	P21	P22	P23
Response Time (milliseconds)	Mean	350.4	323.3	422.6	368.4	426.3	331.6	696.5	310.3
	S.D.	106.0	141.7	249.4	195.1	353.8	121.1	724.4	98.9
Ambient Light (lux)	Mean	332.9	775.7	94.0	101.1	62.7	145.6	177.4	97.0
	S.D.	1191.8	1513.1	281.3	174.5	212.2	271.7	819.7	177.8

Table 2. Mean and standard deviation of PVT response time and ambient light intensity in StudyII.

Person-dependent Model
The person-dependent models achieved a mean RMSE of 36.63 ms (*Max*=110.8, *Min*=17.83, *S.D.*=31.34) and 36.31 ms (*Max*=110.79, *Min*=17.35, *S.D.*=31.23) with the full feature set and the reduced feature set respectively, using 10-fold cross validaton. In the leave-one-day-out cross validation, the mean RMSE was 43.28 ms (*Max*=118, *Min*=17.14, *S.D.*=40.87) and 44.91 ms (*Max*=123.94, *Min*=12.27, *S.D.*=44.01) for models trained using the full feature set and the reduced feature set respectively. The RMSEs are smaller compared to the RMSEs in Study I, which might be attributed to the fact that the high-resolution eye images make pupils easier to detect and the burst images help reduce the effect of bad lighting and angles.

Person-independent Model
The mean RMSE of the person-independent models using leave-one-person-out cross validation was 66.87 ms (*Max*=220.03, *Min*=21.14, *S.D.*=63.68) and 61.33 ms (*Max*=221.09, *Min*=18.41, *S.D.*=65.68) with the full feature set and the reduced feature set respectively. Similar to what we found in Study I, the mean RMSE of person-independent models was higher than person-dependent models as a result of greater variations in pupillary response across different participants. That said, the results from our second study still suggest that passively captured pupil information and ambient light intensity can be used to model users' alertness.

Discussion
In Study II, we addressed some limitations of Study I by using a smartphone with a high-resolution front-facing camera (13MP) and by automatically collecting a burst of pictures whenever participants unlocked their phones. In this way, we could evaluate if our method could be used in a completely passive and unobtrusive way.

As in study I, the results of Study II provide evidence that our method can be used to measure alertness in-the-wild. In addition, our method proved to be more accurate in this study, with lower RMSE scores than the ones found in Study I. The mean RMSE of the predicted MRT using 10-fold cross validation was 36.63 ms using 4 features and 36.31 using only 2 features. The mean RMSE of the generalized model was 66.87 ms using 4 features and 61.33 ms using 2 features.

The fact that the results in Study II were even better than the results of Study I shows that the two changes introduced in Study II (high resolution camera and burst of pictures collected automatically) were effective in increasing the performance of the models. Furthermore, since it was not necessary to make any physical changes in the participants' smartphones, the results show that the method can be easily deployed, requiring only the installation of the mobile application by the user.

GENERAL DISCUSSION
Our alertness level fluctuates over time, and it changes based on several factors, such as circadian rhythms, sleep duration, and stimulant intake. Previous studies have shown different methods to assess alertness, including questionnaires, reaction time test, physiological signals, etc. Recently, researchers proposed methods that can even allow prediction of cognitive states in-the-wild by passively collecting data from smartphones. In this paper, we presented a complementary solution called AlertnessScanner, a computer-vision-based system that models alertness in-the-wild by extracting pupil-to-iris ratio from pictures taken by smartphones, and combining this with other features collected passively. It has the benefits of being unobtrusive as well as not requiring phone use data.

We conducted two in-the-wild studies to evaluate the performance of our solution. In both studies, we evaluated personalized and generalized models using either 4 features (PIR, light, epoch, hours of sleep) or 2 features (PIR, light). Both feature sets can be collected passively from smartphones. We compared the results of our models to the response time obtained using a mobile version of PVT. Response time has been widely used as an indicator of level of alertness [28]. Response time increases as user's level of alertness decreases. Our approach achieved a mean RMSE of 43.28 ms and 66.87 ms for participants' median response time with person-depend model and person-independent model respectively. The results suggest that AlertnessScanner performs better than the method used in the previous study focusing on passively measuring alertness using smartphone usage [5]. As such, our unobtrusive method can serve as a good alternative means of measuring alertness in-the-wild, and potentially enables longitudinal monitoring of alertness in a scalable way.

In the following sections, we discuss how our solution could be used to understand how alertness changes over time, and how knowledge of these alertness changes can be used for real-world interventions.

Understanding Alertness Changes

The fact that AlertnessScanner can be used to measure alertness in-the-wild creates new opportunities for measuring how alertness changes over time. Although our alertness is constantly changing, we are often not aware of these changes and how they affect our daily activities. Having a deeper understanding of our alertness fluctuations may allow us to understand the factors that lead to alertness dips, or in which contexts we are most likely to reach a higher level of alertness. This information can be useful for understand our potential strengths and limitations. For instance, "early rising" is encouraged, and "late rising" is associated with laziness [4]. However, research shows that although some individuals are more likely to be alert early in the morning, others may be more alert in the afternoon or even at night [4]. Being aware of our own alertness patterns allows us to optimize our activities other than following the dictates of society.

In this paper, we decided to focus only on features that could be collected passively from smartphones. In order to have a greater understanding of our alertness changes and the behavioral and contextual factors that affect them, AlertnessScanner could be used in conjunction with existing systems that collect behavioral and environmental data. For example, our method could extend and complement the system developed by Abdullah et al. [5], which enables the collection of other information associated with alertness levels.

Acting upon Alertness Changes

Understanding our alertness cycles can help us come up with long-term solutions for better management of daily activities. For example, our method can allow better scheduling over the day depending on an individual's alertness state and task priorities. Activities that require high alertness, such as writing, could be scheduled to happen during our peak of alertness, while moments of low alertness could be used for resting or for activities that do not require much attention, such as rote tasks [31]. Furthermore, a system based on AlertnessScanner could give personalized recommendations, such as changing sleep habits or avoiding stimulant drinks, in order to manage our alertness level during the day.

A system based on AlertnessScanner can also play an important role in reducing the risk of fatigue related accidents in the workplace. Fatigue is a serious issue where workplace safety is concerned. For example, fatigue has been associated with a 36% increase in serious medical errors [52]. As clinicians increasingly adopt mobile phones and tablets [40, 43], AlertnessScanner can be used to prevent fatigue related accidents in this domain.

Limitations and Future Work

Although our studies provide evidence that our solution can effectively measure alertness in-the-wild, both studies have some limitations.

In Study I, we had to remove the IR filter of the smartphone cameras used by the participants, which is impractical since it compromises other pictures that users take with their phones. Furthermore, in Study I participants had to manually take pictures of themselves using our app, which requires active participation from users.

In Study II, we tried to address the limitations of Study I using front-facing camera with much higher resolution (13MP) and by collecting a burst of pictures passively when users unlocked their phones. This method proved to be unobtrusive, since it did not prompt users for any additional action for data collection. Although the results were better than the ones found in Study I, we acknowledge that currently there are few smartphones on the market with a front-facing camera with the same or better resolution than the one we used. However, as the cameras of smartphones keep improving, we believe that most smartphones will reach this standard eventually.

Another limitation of our work is that although AlertnessScanner takes ambient light into consideration, there is still work to be done to handle very dark or very bright lighting conditions. The quality of images taken in the dark environment can be quite poor. For these images, the segmentation of the iris and pupil is computationally difficult. We can address this issue by augmenting image capture with an additional infrared light source. Since infrared light is invisible, it will not interrupt the work flow of the user and hence would not add to user burden. Moreover, as more and more commercial smartphones are equipped with infrared emitters and infrared cameras for identity authentication, it could help circumvent the challenge of extreme lighting conditions. Environments with very bright light pose similar challenges in terms of accuracy. Under such conditions, the pupil size mostly depends on the brightness of light. However, the embedded light sensors in smartphones can be used to filter out such cases. Moreover, there is recent work in modeling the relationship between light exposure and pupil size [49], which may be applicable to this problem. In future work, we plan to take these models into consideration for better accuracy in bright environments. Apart from extreme lighting conditions, for users who wear eyeglasses, the glare on the eyeglasses also affects the iris and pupil recognition rate in our algorithm. This challenge can be potentially addressed using reflection removal algorithm proposed by Shih et al [44].

In addition, pupil size can also change due to a number of reasons other than alertness states including age [53] and mind-wandering (lost in thoughts) [13] for instance. In our studies, the participants had relatively limited age range. In order to factor in the influence of age, it will be useful to collect and compare data from study population with more diverse age range. And to differentiate mind-wandering from declined alertness, future work should incorporate additional sources of contextual information. For example, blinking and eye gaze [30, 42] as well as patterns of phone use [5] can complement the outcomes of our proposed pupil based algorithms.

Finally, since our system uses facial images to infer alertness, there are some potential privacy concerns. However, it should be noted that our framework does not necessarily need to store images for inferring alertness. Currently, our system computes the PIR offline, but the system can be modified to calculate the PIR from a captured image on-the-fly. Therefore, it is possible to implement the whole data collection process, image capturing, face detection, pupil detection, and PIR calculation within the phone. This would address privacy issues related to facial image capturing.

CONCLUSION

Reduced alertness has adverse effects on our attention, productivity, and even our safety. As such, being able to monitor alertness continuously can have significant impacts across a number of domains. Toward this goal, we developed AlertnessScanner, a computer-vision-based system that allows modeling alertness by extracting pupil-to-iris ratio from pictures taken by smartphones, and combining with other features collected passively. Based on results from two in-the-wild studies, we found that our system is accurate in assessing alertness states, and it can be used to replace methods that are inconvenient to use in-the-wild, such as PVT. Our developed method is a significant step toward granular and continuous assessment of alertness across a large population and over a long period of time. Furthermore, our findings have broader implications for a wide range of applications that can act upon our alertness state.

ACKNOWLEDGMENT

We thank Intel and the Semiconductor Research Corporation for partially supporting the work through the Circadian-Rhythm Aware Computing grant.

REFERENCES

1. 2014. Psychomotor Vigilance Task. `https://en.wikipedia.org/wiki/Psychomotor_vigilance_task`. (2014).

2. 2017a. Cascade Classifier. `http://docs.opencv.org/2.4/doc/tutorials/objdetect/cascade_classifier/cascade_classifier.html`. (2017).

3. 2017b. Morphology. `http://docs.opencv.org/2.4/doc/tutorials/imgproc/opening_closing_hats/opening_closing_hats.html`. (2017).

4. Saeed Abdullah, Mark Matthews, Elizabeth L Murnane, Geri Gay, and Tanzeem Choudhury. 2014. Towards circadian computing: early to bed and early to rise makes some of us unhealthy and sleep deprived. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 673–684.

5. 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. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 178–189.

6. Maysoon F Abulkhair, Hesham A Salman, and Lamiaa F Ibrahim. 2015. Using Mobile Platform to Detect and Alerts Driver Fatigue. *International Journal of Computer Applications* 123, 8 (2015).

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

8. Pedro Arezes, João Santos Baptista, Monica P Barroso, Paula Carneiro, Patrício Cordeiro, Nelson Costa, Rui B Melo, A Sergio Miguel, and Gonçalo Perestrelo. 2013. *Occupational Safety and Hygiene*. CRC Press.

9. Harvey Babkoff, Tamir Caspy, and Mario Mikulincer. 1991. Subjective sleepiness ratings: the effects of sleep deprivation, circadian rhythmicity and cognitive performance. *Sleep: Journal of Sleep Research & Sleep Medicine* (1991).

10. Per Bækgaard, Michael Kai Petersen, and Jakob Eg Larsen. 2015. Assessing Levels of Attention using Low Cost Eye Tracking. *arXiv preprint arXiv:1512.05497* (2015).

11. Joseph V Baranski, Ross A Pigeau, and Robert G Angus. 1994. On the ability to self-monitor cognitive performance during sleep deprivation: a calibration study. *Journal of Sleep Research* 3, 1 (1994), 36–44.

12. Panos Bitsios, Sophia E Schiza, Stella G Giakoumaki, Kyriaki Savidou, Athanasios K Alegakis, and Nikolaos Siafakas. 2006. Pupil miosis within 5 minutes in darkness is a valid and sensitive quantitative measure of alertness: application in daytime sleepiness associated with sleep apnea. *SLEEP-NEW YORK THEN WESTCHESTER-* 29, 11 (2006), 1482.

13. Claire Braboszcz and Arnaud Delorme. 2011. Lost in thoughts: neural markers of low alertness during mind wandering. *Neuroimage* 54, 4 (2011), 3040–3047.

14. Julie Carrier and Timothy H Monk. 2000. Circadian rhythms of performance: new trends. *Chronobiology international* 17, 6 (2000), 719–732.

15. Neil Charness, Michael Tuffiash, Ralf Krampe, Eyal Reingold, and Ekaterina Vasyukova. 2005. The role of deliberate practice in chess expertise. *Applied Cognitive Psychology* 19, 2 (2005), 151–165.

16. Zhenyu Chen, Mu Lin, Fanglin Chen, Nicholas D Lane, Giuseppe Cardone, Rui Wang, Tianxing Li, Yiqiang Chen, Tanzeem Choudhury, and Andrew T Campbell. 2013. Unobtrusive sleep monitoring using smartphones. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*. IEEE, 145–152.

17. John G Daugman. 1993. High confidence visual recognition of persons by a test of statistical independence. *IEEE transactions on pattern analysis and machine intelligence* 15, 11 (1993), 1148–1161.

18. 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.

19. Tilman Dingler, Albrecht Schmidt, and Tonja Machulla. 2017. Building Cognition-Aware Systems: A Mobile Toolkit for Extracting Time-of-Day Fluctuations of Cognitive Performance. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 47.

20. Jill Dorrian, Nicole Lamond, and Drew Dawson. 2000. The ability to self-monitor performance when fatigued. *Journal of Sleep Research* 9, 2 (2000), 137–144.

21. Brice Faraut, Karim Zouaoui Boudjeltia, Michal Dyzma, Alexandre Rousseau, Elodie David, Patricia Stenuit, Thierry Franck, Pierre Van Antwerpen, Michel Vanhaeverbeek, and Myriam Kerkhofs. 2011. Benefits of napping and an extended duration of recovery sleep on alertness and immune cells after acute sleep restriction. *Brain, behavior, and immunity* 25, 1 (2011), 16–24.

22. Danielle J Frey, Pietro Badia, and Kenneth P Wright. 2004. Inter-and intra-individual variability in performance near the circadian nadir during sleep deprivation. *Journal of sleep research* 13, 4 (2004), 305–315.

23. J He, S Roberson, B Fields, J Peng, S Cielocha, and J Coltea. 2013. Fatigue detection using smartphones. *Journal of Ergonomics* 2013 (2013).

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

25. MW Johons. 1991. A new method for measuring daytime sleepiness: The Epworth sleeping scale. *Sleep* 14, 6 (1991), 540–545.

26. Tzyy-Ping Jung, Scott Makeig, Magnus Stensmo, and Terrence J Sejnowski. 1997. Estimating alertness from the EEG power spectrum. *IEEE transactions on biomedical engineering* 44, 1 (1997), 60–69.

27. 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 *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*. IEEE, 248–251.

28. Sylvia Loh, Nicole Lamond, Jill Dorrian, Gregory Roach, and Drew Dawson. 2004. The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers* 36, 2 (2004), 339–346.

29. Otto Lowenstein, Richard Feinberg, and Irene E Loewenfeld. 1963. Pupillary movements during acute and chronic fatigue a new test for the objective evaluation of tiredness. *Investigative Ophthalmology & Visual Science* 2, 2 (1963), 138–157.

30. Jia-Xin Ma, Li-Chen Shi, and Bao-Liang Lu. 2014. An EOG-based Vigilance Estimation Method Applied for Driver Fatigue Detection. *Neuroscience and Biomedical Engineering* 2, 1 (2014), 41–51.

31. Gloria Mark, Shamsi T Iqbal, Mary Czerwinski, and Paul Johns. 2014. Bored mondays and focused afternoons: the rhythm of attention and online activity in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3025–3034.

32. Julio Martín-Moro, Belén de la Fuente, and Fernando Sanz. 2015. Conceptual advantages of using pupil to limbus ratio over absolute pupil diameter. *Journal of neurosciences in rural practice* 6, 4 (2015).

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

34. Hélio Amante Miot, Daniel Rosito Pivotto, Edson Nassib Jorge, and Gláucia Maria Ferreira da Silva Mazeto. 2008. Evaluation of oculometric parameters by facial digital photography: use of iris diameter as a reference. *Arquivos brasileiros de oftalmologia* 71, 5 (2008), 679–683.

35. Deb Kumar Mojumder, Saumil Patel, Kenneth Nugent, John Detoledo, Jongyeol Kim, Nabeel Dar, Henrik Wilms, and others. 2015. Pupil to limbus ratio: Introducing a simple objective measure using two-box method for measuring early anisocoria and progress of pupillary change in the ICU. *Journal of neurosciences in rural practice* 6, 2 (2015), 208.

36. Junpei Nishiyama, Koji Tanida, Masashi Kusumi, and Yutaka Hirata. 2007. The pupil as a possible premonitor of drowsiness. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 1586–1589.

37. Martin Pielot, Tilman Dingler, Jose San Pedro, and Nuria Oliver. 2015. When attention is not scarce-detecting boredom from mobile phone usage. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 825–836.

38. Srinivasan Rajaraman, Sridhar Ramakrishnan, David Thorsley, Nancy J Wesensten, Thomas J Balkin, and Jaques Reifman. 2012. A new metric for quantifying performance impairment on the psychomotor vigilance test. *Journal of sleep research* 21, 6 (2012), 659–674.

39. Riya Regi, Kumar Sai Sailesh, and JK Mukkadan. 2017. Pupil to limbus diameter ratio (pld ratio), olfactory sensitivity and cognition in females with type 2 diabetes mellitus: A case control study. *Research Journal of Pharmacy and Technology* 10, 6 (2017), 1703.

40. Tim Robinson, Thomas Cronin, Haider Ibrahim, Mark Jinks, Timothy Molitor, Joshua Newman, and Jonathan Shapiro. 2013. Smartphone use and acceptability among clinical medical students: a questionnaire-based study. *Journal of medical systems* 37, 3 (2013), 1–7.

41. ARUNA SAJEEVAN and KUMAR SAI SAILESH. 2017. CORRELATION OF PUPIL TO LIMBUS DIAMETER RATIO (PLD RATIO) WITH BLOOD PRESSURE AND PULSE RATE. *Int J Pharma Bio Sci* 8, 3 (2017), 12–16.

42. Robert Schleicher, Niels Galley, Susanne Briest, and Lars Galley. 2008. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics* 51, 7 (2008), 982–1010.

43. Joseph Sclafani, Timothy F Tirrell, and Orrin I Franko. 2013. Mobile tablet use among academic physicians and trainees. *Journal of medical systems* 37, 1 (2013), 1–6.

44. YiChang Shih, Dilip Krishnan, Fredo Durand, and William T Freeman. 2015. Reflection removal using ghosting cues. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3193–3201.

45. Maria Thomas, Helen Sing, Gregory Belenky, Henry Holcomb, Helen Mayberg, Robert Dannals, JR Wagner, David Thorne, Kathryn Popp, Laura Rowland, and others. 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.

46. HPA Van Dongen, Maurice D Baynard, Greg Maislin, and David F Dinges. 2004. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *SLEEP-NEW YORK THEN WESTCHESTER-* 27, 3 (2004), 423–433.

47. Antoine U Viola, Lynette M James, Luc JM Schlangen, and Derk-Jan Dijk. 2008. Blue-enriched white light in the workplace improves self-reported alertness, performance and sleep quality. *Scandinavian journal of work, environment & health* (2008), 297–306.

48. Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 3–14.

49. Andrew B Watson and John I Yellott. 2012. A unified formula for light-adapted pupil size. *Journal of vision* 12, 10 (2012), 12–12.

50. Sophie MT Wehrens, Shelagh M Hampton, Myriam Kerkhofs, and Debra J Skene. 2012. Mood, alertness, and performance in response to sleep deprivation and recovery sleep in experienced shiftworkers versus non-shiftworkers. *Chronobiology International* 29, 5 (2012), 537–548.

51. Barbara Wilhelm, Helmut Wilhelm, H Lüdtke, Peter Streicher, and Martin Adler. 1998. Pupillographic assessment of sleepiness in sleep-deprived healthy subjects. *Sleep* 21, 3 (1998), 258–265.

52. Ann Williamson, David A Lombardi, Simon Folkard, Jane Stutts, Theodore K Courtney, and Jennie L Connor. 2011. The link between fatigue and safety. *Accident Analysis & Prevention* 43, 2 (2011), 498–515.

53. Barry Winn, David Whitaker, David B Elliott, and Nicholas J Phillips. 1994. Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science* 35, 3 (1994), 1132–1137.

54. Robert E Yoss, Norma J Moyer, and Robert W Hollenhorst. 1970. Pupil size and spontaneous pupillary waves associated with alertness, drowsiness, and sleep. *Neurology* 20, 6 (1970), 545–545.