2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Verification methodology for Smart Awakening Systems

Denys Sverdlov[1], Valerii Dziubliuk[2], Kostyantyn Slyusarenko[3], Yevhen Romaniak[4] and Anastasiia Smielova[5]

Abstract— **A mental and physical recovery after an awakening moment depends not only on the overall sleep duration and quality but mostly on the sleep stage in the waking moment. The most comfortable awakening moment is during the Light or Wake sleep stages. But the fix-time alarm clock doesn't take into account the sleep stage in the awakening moment, which often results in awakening during the Deep or Rapid Eyes Movement stages. To reduce the negative recovery effect, big companies and research groups develop various awakening systems. Such systems recognize sleep stages based on wearable sensors' data (mostly from accelerometer sensors) and thus can find the easiest awakening moment time with minimal recovery effects.**

However, it is quite hard to measure and verify the efficiency of such systems without using polysomnography (which can be performed only in clinical conditions by medical experts). To solve this problem we developed a methodology based on questionnaires and psychological tests. Such an approach has big scalability, does not require special medical equipment, and can be evaluated in a home environment with minimal support effort. The proposed verification approach has been tested on smartwatches with the sleep stages forecast model. The proposed model accuracy was 78%. Results of our experiment show that the majority of users demonstrated a correlation between awakening quality and the verification tests performance.

I. INTRODUCTION

Sleep plays an important role in the human lifetime. It provides physiological and psychological relaxation, which affects many aspects of health and quality of life: daily activity, physical disorders, cardiovascular diseases, diabetes, obesity, etc [1]. Physiologically, sleep is a cyclical process of changing four sleep stages: Rapid Eye Movement (REM), Light, Deep, and Wake stages [2].

Sleep quality is determined by the cyclical pattern of successive sleep stages changing. The ideal night sleep pattern includes 4–6 completed sleep cycles of Light-Deep-Light-REM stages [3]. But human life often requires waking in specific moments, usually early in the morning. Such sleep interruption instead of natural awakening may result in an uncomfortable state after awakening, called sleep inertia. Sleep inertia is a physiological state of impaired cognitive and sensory-motor performance that is present immediately after waking, when a person experiences drowsiness, disorientation and a decline in motor dexterity [4], [5]. Impairment from sleep inertia may take several hours to dissipate. Usually, morning sleep inertia is experienced for 15 to 60 minutes after waking [6].

Usually, awakening in the Deep/REM sleep stages is associated with greater sleep inertia [7], [8]. Generally, awakening in the Light/Wake sleep stages is much more comfortable. To find the optimal moment of awakening, an ideal alarm system should take into account current and forecast future sleep stages to find the optimal moment for awakening (in the Wake or Light sleep stages). An example of such a system based on wearable sensor data from photoplethysmography (PPG) and accelerometer (ACC) data was proposed in [3].

The development of such systems is complicated since the verification of the effectiveness of such algorithms is carried out with the involvement of specialized medical equipment, and medical personnel, for conducting polysomnography (PSG). The gold standard of sleep stages classification, based on the PSG approach, includes recording electrocardiogram, electrooculogram, body posture monitoring, nasal pressure, oronasal airflow, thoracic and abdominal volume changes, snoring sound and blood oxygen saturation [3]. These methods make it possible to achieve sleep stages classification accuracy of about 0.9 for 5-sleep stages classes [9], [10] and 0.80–0.95 for sleep/wake classification [1], [11], [12]. However, PSG measurements have certain limitations: require medical sleep experts and equipment for sleep stages classification.

Based on the alarm system proposed in the work [3], in the current work, we built our model which was able to distinguish the Wake and Light from the Deep and REM sleep stages for predicting future sleep stages to find the global optimum. In the forecast model, we embedded additional input heads which process independent different time slices of input signals (from one to sixty seconds). Such modification gave 78% sleep stages forecast accuracy for Light/Wake vs. Deep/REM classification. To verify this model, we proposed a method for measuring awakening quality by conducting psychological tests and questionnaires. Such methodologies are often used to assess the full night sleep quality and efficiency of person's sleep [6], [8], [13], [14], [15]. But we did not find any studies that propose methods for detection waking up quality and post-sleep inertia. So in the present article, we propose a verification method based on the hypothesis, that user's feelings and mental cognitive abilities immediately after awakening in the Light/Wake sleep phases are higher in comparison with

[1]Denys Sverdlov is with Samsung R&D Institute Ukraine, Kyiv, 01032, Ukraine d.sverdlov@samsung.com
[2]Valerii Dziubliuk is with Samsung R&D Institute Ukraine, Kyiv, 01032, Ukraine v.dzyublyuk@samsung.com
[3]Kostyantyn Slyusarenko is with Samsung R&D Institute Ukraine, Kyiv, 01032, Ukraine k.slyusarenk@samsung.com
[4]Yevhen Romaniak is with Samsung R&D Institute Ukraine, Kyiv, 01032, Ukraine i.romaniak@samsung.com
[5]Anastasiia Smielova is with Samsung R&D Institute Ukraine, Kyiv, 01032, Ukraine a.smielova@samsung.com
*This work is supported by Samsung R&D Institute Ukraine

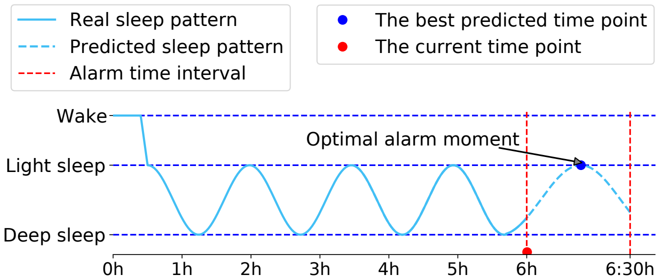

Fig. 1. An example of the optimal alarm moment prediction. The dashed blue curve represents the predicted sleep pattern of a user (represented by the probability of Light/Wake classes.

those after awakening in the Deep/REM sleep phase. Such methodology shows that 78% accuracy of the sleep stages forecast model was enough to improve awaking quality for most of the users.

The article is organized as follows. In Section II we give a short description of the architecture of the proposed Smart Awakening System introduced in [3]. In Section III we describe our verification methodology and metrics we used to measure the efficiency of our system. In Section IV, we introduce the experimental database of users of our awaking system. Finally, in Section V we describe and analyze users' answers and performances in psychological tests according to our verification methodology.

II. Smart Awakening System Architecture

The overview of the proposed Smart Awakening System architecture is shown in Fig. 2. The core of technology was 2-classes sleep stages forecasting model for the 20-min interval. It required 200 minutes of consecutive sleep data from PPG and ACC sensors with a frequency of 20 Hz. The main difference from the model presented in work [3] was the input signal was split into non-overlapping N-seconds windows. Each data split was fed in a separate rule-based feature extractor. The rule-based features extractor performed signal statistic calculation in the time and frequency domain on the slide window (mean, standard deviation, median, maximum, minimum, and quantile). Only 60 seconds window size features were calculated similarly as it was presented in [3]. These rule based features were fed to the trainable convolution neural network (CNN) features extraction model. For each N-second window size we built a separate CNN model. The CNN model was built in such a way that output features of length 200 with 64 channels for all N-second windows. As building blocks for the model MaxPool, BatchNorm, Dropout, and CNN layers were used with kernel size 3 or 5 and fixed number of 32 channels. To forecast sleep stages, the CNN features were concatenated in one feature tensor which was fed to Gated Recurrent Units sequential model (GRU) [16]. The output of the forecast model consisted of 20 minutes forecast probabilities of the Light/Wake sleep stages. At the end number of trainable parameters was 463786. Probabilities were used to calculate the optimal time for an alarm clock signal for waking up

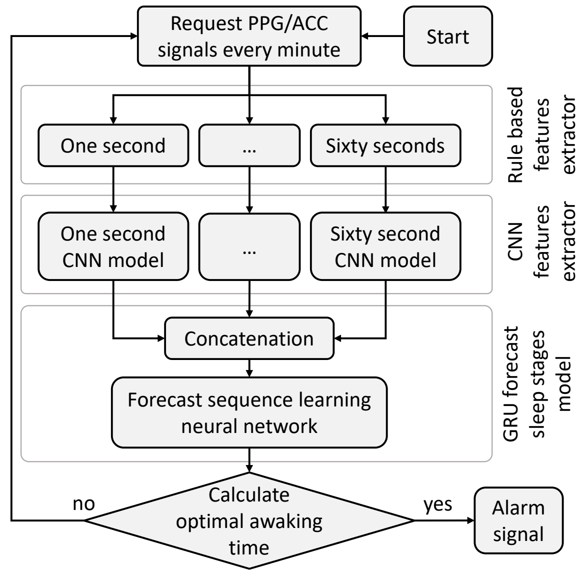

Fig. 2. The overview of proposed Smart Awakening System architecture.

a user in the most comfortable time moment (inside time period defined by the user).

The system recalculates the optimal alarm moment every minute till the current moment becomes optimal. At that moment an alarm clock goes off.

The Smart Awakening System was implemented as a pair of smartwatches and smartphones.

III. Description of Verification Methodology

As was mention before, the state-of-the-art method for full night sleep stages monitoring and detecting the awakening sleep stage is based on polysomnography. But this approach is difficult to perform in non-medical conditions and the data is disturbed by the presence of medical staff and numerous medical equipment, so call "white coat syndrome" [17]. This reason makes it hard to understand the efficiency and benefits of developing new approaches for improving sleep and awakening quality. To solve this problem we proposed the method of awaking quality validation.

Waking up quality is measured in two ways:

- **Subjective** user wake up rate using **a questionnaire** for reflecting personal opinion. We made 2 questionnaires: Waking quiz (performed immediately after user awakening) and Get up quiz (at least after 30 minutes since awakening). The goal of the Waking quiz was to measure the first subjective impression of the quality of the awakening by the Smart alarm application. Get up quiz questions were selected from specialized sleep quality questionnaires [15], [18], [19]. The main goal of this quiz was to find untypical users and unusual sleep conditions during the verification procedure and measure the subjective impression of awakening quality after some time after getting out of a bed. In this quiz, the user answered the questions about the condition before sleep, during sleep, and waking condition.

- **Psychological tests** for reflecting physiological user's state and cognitive mental abilities after awaking. We chose five psychological tests: Stroop, Finger Tapping, Sound Alertness, Multitasking, N-back tests. These tests are widely used in psychology to assess different aspects of user mental health and cognitive abilities: working memory, concentration, motor speed, selective attention, automaticity, parallel distributed processing, executive functions, and choice reaction time.

A. Psychological Tests Implementation

All tests were implemented as an Android application on a smartphone. Gamification of tests was done with a slightly modified tests procedure and shortened test time duration. Such a solution allowed to perform verification of awaking quality more easily and quickly. Users didn't need to change their environment, regime (go to sleep, get up), and habits for alarm clock verification procedure.

1) Stroop Test (processing speed, selective attention, automaticity, parallel distributed processing [13]): The test is performed with four colors: Red (R), Green (G), Blue (B), and Yellow (Y). A user should choose the button from four options "R", "G", "B", and "Y" which correspond to the ink color of the color caption, and press it as fast as possible.

2) Finger Tapping Test (motor speed [20]): A user should tap as fast as possible for 10 seconds on the phone's screen to achieve the highest amount of taps.

3) Sound Alertness Test (choice reaction time [14]): A user should respond to sound commands: "Left", "Right", "Up", "Down"; by swiping as fast as possible to the corresponding direction. The whole test consists of a sequence of 15 aforementioned sound commands in random order.

4) Multitasking Test (executive functions, processing speed, selective attention, automaticity, parallel distributed processing [21]): The task includes 2 subtasks: "Filling" and "Shape", which are selected randomly, and a user should reply to each of them. In the "Shape" subtask a user should classify the shape of the presented figure between 'diamond' (by pressing button "L") or 'rectangle' (button "R"). In the "Filling" subtask a user should classify the filling of the presented figure: 'two dots' (by pressing button "L") or 'three dots' (button "R"). In each trial random subtask mode is written, the figure has a random shape ('diamond' or 'rectangle') and an amount of dots (2 or 3). A user should press the right button ("R" or "L") for the current subtask ignoring distraction: in the "Filling" subtask figure's shape is not important.

5) N-back Test (working memory and concentration [22], [23]): In the N-back (in our case – 3-back) test one letter appears in the center of the screen, and a user should respond whether the current letter matches the letter presented three trials previously as fast as possible. If a user decides that it is the same letter, he/she should press the "M" button, in the opposite case – the button "N". A user can respond to stimuli immediately after letter appearance. The test ends when 25 letters were provided for remembering and recalling.

In all tests, if a user doesn't respond to a task for 3 seconds, the algorithm considers this case as failed and shows the next task. The duration of Multitasking, Stroop tests is 60 seconds each. For Multitasking, Stroop tests approximately half of the tasks are congruent: stimuli and distractor mean the same and requires pressing the same button (a figure with a 'diamond' shape and 2 dots in the Multitasking task requires pressing "L" whether it is the "Filling" or the "Shape" subtask).

B. Smart Awakening System in Verification Mode

To measure the sleep inertia, a user wakes up in different conditions:

- The "Best" mode is the regular regime of alarm clock which wakes up a user in the Light/Wake stage.
- The "Worst" mode is a specially modified alarm system application for waking up a user in the Deep/REM stage.
- The "Natural" awakening mode in a case when a user wakes up by himself/herself.
- Cases of awakening due to external distraction were not taken into consideration.

During testing, the user doesn't know in which mode (the "Best" or the "Worst") the alarm is working. The application automatically and randomly switches between the "Best" and the "Worst" modes. Measuring and analyzing differences in user performance between waking up psychological tests and Waking quiz responses will show how the alarm application can influence the quality of awakening. It is expected to get higher subjective awakening quality and better performance in psychological tests in the "Best" and "Natural" cases in comparison with those in the "Worst" case.

The pipeline for one trial is following:

- Before a sleeping user sets an alarm time period.
- After awaking user answers the Waking quiz and physiological tests.
- After fully waking up (at least after 30 minutes) user answers Get up quiz the tests.

This procedure was repeated for each user until 7 trials in each mode ("Best", "Worst", "Natural") were collected. A few first trials were for adaptation to the tests and evaluation procedure. During the verification, a user can wake up before the alarm ringing. In this case, a user can cancel the alarm clock and start the Waking quiz and tests. In case of external distraction awakening, it is not needed to complete tests and quizzes, because such cases were not evaluated. To reduce the influence of body pose, at the start of testing, a user has to choose a body pose (lying or sitting, usage of one or two hands, etc.), and pass physiological tests only in it.

C. Metrics

User Satisfaction Index allows describing the amount of users in the experiment who confirm better well-being when awaking in the optimal time estimated by our model. It is based on the comparison of user awakening rates in different application modes ("Best" and "Worst"):

$$UD_k = \frac{1}{N_{bk}} \sum_{i=1}^{N_{bk}} r_i^{best} - \frac{1}{N_{wk}} \sum_{i=1}^{N_{wk}} r_i^{worst}, \quad (1)$$

where UD_k – the differentiation rate for user k (the difference between self-feeling during awakening in "Best" and "Worst" modes given in Wake up quiz), N_{bk}, N_{wk} – the number of filtered responses in "Best" and "Worst" awakening modes for user k, r_i – the response rate of user i filtered by his/her daily lifestyle and usual sleep conditions (in "Best" or "Worst" mode). The rate is the user's subjective evaluation of awakening quality. It ranges between 1 and 9, with the step of 0.1.

The Satisfaction Index (SI) is calculated as the average positive differentiation rate among all participants:

$$SI = \frac{1}{N} \sum_{k=1}^{N} I[UD_k > 0], \quad (2)$$

where $I[\cdot]$ – the indicator function, N – the overall number of participants. In the ideal case $SI \to 1$ (all user's awakening rates during the "Best" mode are higher than in the "Worst" mode).

Performance in Psychological Tests: We analyzed user's performance in psychological tests using two metrics: number of mistakes and reaction speed (in ms). For Stroop Test, there is an additional metric called the Stroop effect which is computed as a difference between reaction speed in trials with corresponding and different color and text which represents this color. As for Tap Count Test, we refer only to the number of taps made by a user.

In this study, our hypotheses about performance in the tests are following:

- Mistakes and Reaction in all tests: it is expected to see less number of mistakes and better reaction time during a psychological test in the "Best" alarm mode.
- Stroop Effect Value in Stroop Effect Test: it is expected to see a lower value of Stroop effect in the "Best" alarm mode.
- Taps Count in Tap Counter Test: it is expected to see a higher number of taps in the "Best" alarm mode.

IV. EXPERIMENTAL DATA

The sleep stages forecast model was trained on the sleep dataset collected by Samsung Medical Center (SMC) during the 2017 year. The train/test set included data of subjects with Apnea-Hypopnea Index less than 15, consisted of 240 nights from 187 different subjects. Sleep data were labeled using PSG data for every 30s epochs using the American Academy of Sleep Medicine guideline [24]. The data contains wrist green light PPG and 3D–ACC signals from Samsung Galaxy Watch at 20Hz sampling frequency.

The verification of the proposed Smart Awakening System was conducted in the 2021 year from February to April. The study protocol respected the Helsinki declaration and an informed consent was obtained from each participant before starting the experiment. In our study, 9 participants used the application during the test period of 24 days. Thus, result data consists of 217 quizzes and psychological test results.

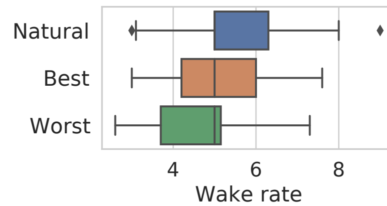

Fig. 3. Distribution of awakening quality rate for Wake up quiz depending on the alarm mode.

V. RESULTS

A. The Sleep Stages Forecast Model

The sleep stages forecast model was trained to classify Light/Wake vs. Deep/REM sleep phase. We used a cross-validation (5% of train subject logs, 5 folds) for model hyperparameters tuning. During the training and test phases, the logs were split on 220 minutes windows (200 is a log history, the last 20 is sleep stages labels to forecast) using sliding window approach with 1 minute stride. All models' weights were initialized with He Normal initialization [25]. Cross Entropy loss function and Adam optimizer with cycle learning rate [26] (ranging from 3×10^{-4} to 3×10^{-7}) were used for each run. The batch size during training was 64. L2 regularization coefficient was 1×10^{-3}. To test the model users from the SMC dataset were randomly split between the train and test parts with a percentage of 0.8:0.2. The best accuracy was obtained for the model with N-seconds windows where N was 1 and 60 seconds. These modifications allowed to achieve 78.0% and 77.5% (vs. 77% and 75% in [3]) accuracy for 10 and 20 minutes forecast respectively.

B. User Satisfaction Index

Satisfaction Index (2) was calculated based on the users' awakening quality rate in the Wake up quiz. The distribution of users' differentiation rates UD_k is shown in Fig. 3. Such overlapping of UD_k rates in different modes can happen due to the users' regular sleep schedule, daily lifestyle, and sleep condition [6], [27]. On average users give higher awakening rates in the "Best" alarm mode than in the "Worst" mode, which indicates better subjective feeling when awakening in the "Best" mode (corresponds to "Light/Wake" sleep phase). This resulted in User Satisfaction Index (2) among all the participants equal to 0.89.

C. Psychological Tests

Distributions of mistakes made by all users in test trials in different awakening states ("Best", "Worst", "Natural" alarm modes) and reaction speed are shown in Fig. 4. To visualize user's performance (mistakes and reaction speed) on the same scale, these values for each user in each trial were normalized by the user's median value in all trials:

$$N_{ik(normalized)} = \frac{N_{ik}}{median([N_{i1}, N_{i2}, ..., N_{in}])}, \quad (3)$$

where N_{ik} – the number of mistakes or reaction speed for user i in trial k.

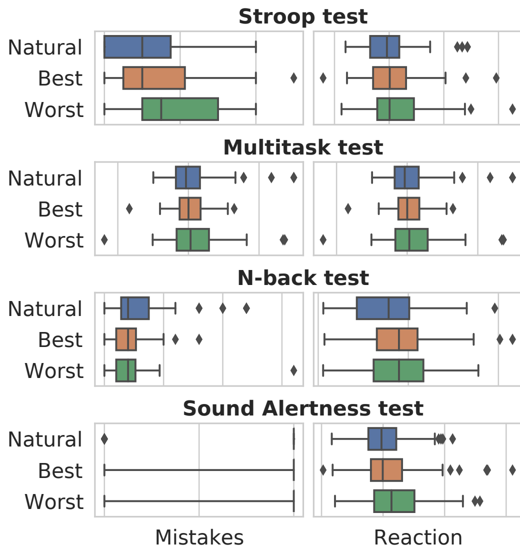

Fig. 4. Distributions of mistakes (left parts) and reactions (right parts) for the psychological tests depending on the alarm mode.

For the Stroop Test, on average users make fewer mistakes in the "Best" alarm mode (corresponds to awakening in the "Light/Wake" sleep phase) than in the "Worst" mode (corresponds to awakening in the "Deep/REM" sleep phase). The least number of mistakes is made after waking without alarm. The Pearson correlation coefficient between the numbers of mistakes and waking rates is equal to -0.31 – on average higher awakening rates correspond to better performance (fewer number of mistakes) in the Stroop Test. As for the reaction speed and the Stroop effect metric, the correlation with the awakening mode is less significant in our study.

User performance in the other psychological tests described in Section III didn't show a significant correlation with the alarm mode in our study (see Fig. 4). There are several factors that may explain these results. First, when passing tests some cognitive abilities, e.g. concentration, memory and attention, are increased. Tests also create stressful situations that might result in Cortisol Awakening Response, which is connected with sleep inertia dissipation [28], [29]. Second, the effect of blue-enhanced light from a mobile display and sounds in tests (Sound Alertness) might increase the speed of transition to a fully awake state [30]. Also, some of the participants have regular sleep schedule and habits, which may decrease sleep inertia severity and duration [6], [27]. These factors potentially may affect the performance during testing immediately after awaking. Clarification of these hypotheses may require additional testing with a bigger number of participants.

VI. CONCLUSIONS

The results of measurements proved the efficiency of our Smart Awakening System. The majority of users reported a better feeling of well-being when awakening in the Light sleep phase. It was shown that Stroop Test, performed immediately after awakening is efficient to measure the sleep inertia effect after awakening and its dependence on different sleep stages in the awakening moment. On the other hand, it was shown that psychological tests which don't perform immediately after awakening were not suitable to measure such effects.

The majority of users reported a better wake-up rate in awakening during the "Best" mode by using our sleep stages forecast model. These results conclude that such Smart Awakening Systems (which achieved at least 78% accuracy) are suitable for daily use to improve users' awakening quality. The future work will focus on sleep latency detection and reduction. Also, additional research is required to develop a psychological test that can combine different aspects of users' mental, physical, and cognitive states with just enough time duration to assess sleep inertia severity. Another important topic is the efficiency investigation of such systems under conditions of sleep deprivation when people can sleep only for short periods and have to perform stressful tasks after (soldiers, pilots, etc.).

REFERENCES

[1] L. Xie, H. Kang, Q. Xu, M. Chen, Y. Liao, M. Thiyagarajan *et al.*, "Sleep drives metabolite clearance from the adult brain," *Science (New York, N.Y.)*, vol. 342, pp. 373–7, 10 2013.

[2] E. Giri, A. Arymurthy, M. I. Fanany, and S. Wijaya, "Sleep stages classification using shallow classifiers," 10 2015.

[3] K. Slyusarenko and I. Fedorin, "Smart alarm based on sleep stages prediction," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 4286–4289.

[4] P. Tassi and A. Muzet, "Sleep inertia," *Sleep Medicine Reviews*, vol. 4, no. 4, pp. 341–353, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1087079200900984

[5] M. Shahin, L. Mulaffer, T. Penzel, and B. Ahmed, "A two stage approach for the automatic detection of insomnia," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2018, p. 466—469, July 2018. [Online]. Available: https://doi.org/10.1109/EMBC.2018.8512360

[6] M. Jewett, J. Wyatt, A. Ritz-De Cecco, S. Khalsa, D. Dijk, and C. Czeisler, "Time course of sleep inertia dissipation in human performance and alertness," *Journal of sleep research*, vol. 8, no. 1, p. 1—8, March 1999. [Online]. Available: https://doi.org/10.1046/j.1365-2869.1999.00001.x

[7] P. Tassi, A. Bonnefond, O. Engasser, A. Hoeft, R. Eschenlauer, and A. Muzet, "Eeg spectral power and cognitive performance during sleep inertia: The effect of normal sleep duration and partial sleep deprivation," *Physiology and behavior*, vol. 87, pp. 177–84, 02 2006.

[8] N. Santhi, A. Lázár, P. McCabe, J. Lo, J. Groeger, and D.-J. Dijk, "Sex differences in the circadian regulation of sleep and waking cognition in humans," *Proceedings of the National Academy of Sciences*, vol. 113, p. 201521637, 04 2016.

[9] R. Rosenberg and S. Hout, "The american academy of sleep medicine inter-scorer reliability program: Sleep stage scoring," *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine*, vol. 9, pp. 81–7, 01 2013.

[10] P. Finan, J. Richards, C. Gamaldo, D. Han, J. Leoutsakos, R. Salas *et al.*, "Validation of a wireless, self-application, ambulatory electroencephalographic sleep monitoring device in healthy volunteers," *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine*, vol. 12, 09 2016.

[11] S. Ancoli-Israel, R. Cole, C. Alessi, M. Chambers, W. Moorcroft, and C. P. Pollak, "The role of actigraphy in the study of sleep and circadian rhythms," *Sleep*, vol. 26, no. 3, p. 342—392, May 2003. [Online]. Available: https://doi.org/10.1093/sleep/26.3.342

[12] A. Hakim, "Wrist actigraphy," *Chest*, vol. 139, pp. 1514–27, 06 2011.

[13] P. Sagaspe, M. Sanchez-Ortuno, A. Charles, J. Taillard, C. Valtat, B. Bioulac *et al.*, "Effects of sleep deprivation on color-word, emotional, and specific stroop interference and on self-reported anxiety," *Brain and cognition*, vol. 60, pp. 76–87, 03 2006.

[14] M. Lamers, A. Roelofs, and I. Rabeling, "Selective attention and response set in the stroop task," *Memory and cognition*, vol. 38, pp. 893–904, 10 2010.

[15] D. J. Buysse, C. F. Reynolds III, T. H. Monk, S. R. Berman, and D. J. Kupfer, "The pittsburgh sleep quality index: a new instrument for psychiatric practice and research," *Psychiatry research*, vol. 28, no. 2, pp. 193–213, 1989.

[16] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk *et al.*, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734. [Online]. Available: https://www.aclweb.org/anthology/D14-1179

[17] Z. Kaya, M. Kayrak, K. Demir, M. Gündüz, H. Alibasiç, M. Karanfil *et al.*, "The relationship between white coat hypertension and sleep quality," *Sleep and Biological Rhythms*, vol. 12, no. 3, pp. 203–211, 2014.

[18] M. W. Johns, "A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale," *Sleep*, vol. 14, no. 6, pp. 540–545, 11 1991. [Online]. Available: https://doi.org/10.1093/sleep/14.6.540

[19] C. Carney, D. Buysse, S. Ancoli-Israel, J. Edinger, A. Krystal, K. Lichstein *et al.*, "The consensus sleep diary: Standardizing prospective sleep self-monitoring," *Sleep*, vol. 35, pp. 287–302, 02 2012.

[20] P. Schatz, "Finger tapping test," *Encyclopedia of Clinical Neuropsychology*, vol. 1, 01 2011.

[21] K. Wilckens, S. Woo, K. Erickson, and M. Wheeler, "Sleep continuity and total sleep time are associated with task-switching and preparation in young and older adults," *Journal of Sleep Research*, vol. 23, 04 2014.

[22] J. Cousins, K. Sasmita, and M. Chee, "Memory encoding is impaired after multiple nights of partial sleep restriction," *Journal of Sleep Research*, vol. 27, 07 2017.

[23] W. Kirchner, "Age differences in short-term retention of rapidly changing information," *Journal of experimental psychology*, vol. 55, no. 4, p. 352—358, April 1958. [Online]. Available: https://doi.org/10.1037/h0043688

[24] R. Berry, R. Brooks, C. Gamaldo, S. Harding, R. Lloyd, S. Quan *et al.*, "The aasm manual for the scoring of sleep and associated events: Rules, terminology and technical specifications, version 2.5," *American Academy of Sleep Medicine*, 2018.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, ser. ICCV '15. USA: IEEE Computer Society, 2015, p. 1026–1034. [Online]. Available: https://doi.org/10.1109/ICCV.2015.123

[26] L. N. Smith, "No more pesky learning rate guessing games," *CoRR*, vol. abs/1506.01186, 2015. [Online]. Available: http://arxiv.org/abs/1506.01186

[27] D. F. Dinges, "Are you awake? cognitive performance and reverie during the hypnopompic state." 1990.

[28] R. Law and A. Clow, "Stress, the cortisol awakening response and cognitive function," *International review of neurobiology*, vol. 150, pp. 187–217, 2020.

[29] S. Hall, B. Aisbett, J. Tait, A. Turner, S. Ferguson, and L. Main, "The acute physiological stress response to an emergency alarm and mobilization during the day and at night," *Noise and Health*, vol. 18, p. 150, 05 2016.

[30] N. Santhi, J. A. Groeger, S. N. Archer, M. Gimenez, L. J. Schlangen, and D.-J. Dijk, "Morning sleep inertia in alertness and performance: effect of cognitive domain and white light conditions," *PloS one*, vol. 8, no. 11, p. e79688, 2013.