

# Exploring Subjective Sleepiness with UnnCyberpsy, a Web-Based Psychophysiological Research Tool


Valeriia Demareva   [ Lobachevsky State University | [valeriia.demareva@fsn.unn.ru](mailto:valeriia.demareva@fsn.unn.ru) ]

Valeriia Viakhireva  [ Lobachevsky State University | [vvv@fsn.unn.ru](mailto:vvv@fsn.unn.ru) ]

Irina Zayceva  [ Lobachevsky State University | [irozy@yandex.ru](mailto:irozy@yandex.ru) ]

Andrey Demarev  [ Lobachevsky State University | [demarev@unn.ru](mailto:demarev@unn.ru) ]

Nikolay Nazarov  [ Lobachevsky State University | [nicola.dzr@gmail.com](mailto:nicola.dzr@gmail.com) ]

 Faculty of Social Sciences, Lobachevsky State University, 7 University Lane, Nizhny Novgorod, 603950, Russia.

Received: 05 December 2023 • Accepted: 08 August 2024 • Published: 01 September 2024

**Abstract** The foundation of scientific research lies in robust data collection methodologies, increasingly embracing a data-driven paradigm within the realm of psychological studies. This paper outlines our extensive undertaking of a psychophysiological investigation utilizing the web-based platform UnnCyberpsy for seamless data aggregation and processing. Data collection stands as a pivotal phase in scientific inquiry, with an increasing inclination toward data-centric methodologies within psychological research. This study delineates our endeavor in conducting a comprehensive psychophysiological investigation employing the web-based tool UnnCyberpsy for data acquisition and analysis. Constructed within the PHP programming framework and ‘CodeIgniter’ microframework (version 4.0), UnnCyberpsy offers pivotal functionalities, including scheduling equipment pick-ups, guiding through test procedures using a branched algorithm, and handling data storage and preprocessing. Each test phase is pre-coded into the system, empowering participants to independently undergo the testing regimen via personal devices. Automatic data collection, preprocessing, and storage streamline the process, ensuring comprehensive participant data validation and preservation upon experiment completion. Embracing UnnCyberpsy resulted in an elimination of data loss risk to 0%, a 30% rise in potential participant engagement, and expedited data gathering constrained solely by participant numbers and equipment availability. Our exploration underscores the pivotal role of online applications like UnnCyberpsy, not only streamlining data collection but also fostering a symbiotic environment benefiting both researchers and participants. This paradigm shift towards digital platforms stands as a beacon for future research endeavors, promising unparalleled convenience and efficiency in the realm of data-driven scientific inquiry.

**Keywords:** Web-Application, Psychophysiological Research, Subjective Sleepiness

## 1 Introduction

The evolution of scientific inquiry is intrinsically linked to the refinement of data collection methodologies, an aspect accentuated by the burgeoning influence of information technologies. Within the dynamic landscape of psychological research, the burgeoning prevalence of data-driven approaches underscores the ongoing quest for automating data collection and processing[?]. This imperative transition stems not only from the need for efficiency but also from the escalating concern about data quality amid the surge in big data analysis. The traditional manual gathering of data, inherently sluggish and predisposed to human errors, confronts profound challenges in meeting the demands of modern research endeavors.

In recent years, the development of intelligent systems for detecting the drowsiness condition of human operators has become crucial for preventing human-caused accidents. Therefore, new research is necessary to develop reliable alerting methods for recognizing episodes of drowsiness during human interaction with complex technical systems. While many studies have focused on developing intelligent warning methods for smart vehicles, achieving up to 97% accuracy [Phan *et al.*, 2023] in preventing accidents caused by driver drowsiness [Jabbar *et al.*, 2018][Mohana and Rani,

2019][Wong and Lau, 2019], there are currently no studies analyzing large amounts of data for reasonably representative samples. Our project stands out from others in that it is based on a representative sample of a specialized population involved in a critical process within a specific territory — namely, drivers. Moreover, we will collect, for the first time, a complexly marked large dataset with multi-hour heart rhythm recordings, marked for sleepiness using three measures: subjective sleepiness according to self-assessment scales, daily hour, and time of the last questionnaire filling.

In this paper we share the experience of conducting mass psychophysiological research using our own web-application UnnCyberpsy for data collection and processing. The scientific goal of our project was to further create sleepiness detector based on the analysis of heart rate data.

To collect big data suitable for searching for markers of different sleepiness levels in heart rate patterns, it is needed to: (a) invite the participant to the laboratory to give him instructions and equipment; (b) get information about the participant’s sleepiness level at different time points to get individual dynamics; (c) get information when a person went to sleep, to understand when the period before falling asleep approximately was; (d) collect data on the participant’s socio-demographic characteristics of in order to identify the influence of various factors and make the detector more accu-

rate. The articles [Bizzego *et al.*, 2014][Thukral and Goel, 2012][Hanbury *et al.*, 2019][Zhang *et al.*, 2017][Canino *et al.*, 2016] indicate the advantages of using automated systems for data collection: high speed of data collection, improving the quality and completeness of data, and reducing the data recipient's workload. Due to the above-mentioned (a)-(d) aspects of project research, and due to the need to conduct research on a relatively large group of people, the highlighted advantages of using automated systems disposed us to the decision to use an automated system to collect all the necessary data. Systems in the form of a web-application (or web-based systems) were preferable as there is no need for installation and local data storage.

## 2 Related works

The challenges researchers face in data gathering have led to the development of diverse software solutions. In a review of available software options by C. Alavi and J. Massman, they compared these options based on features provided and associated costs [Alavi and Massman, 2016]. Electronic data management systems range from simple services to full-scale solutions with administrative support. The researchers categorized available sources into five categories, ranging from stand-alone spreadsheets in Microsoft Excel and simple databases in Microsoft Access to high-cost, large-scale database systems and cloud-based solutions. However, not all these categories are secure enough to contain participants' personal information, despite being accessible and easy to use. On the other hand, some options do not allow for customization of survey forms, while others require extensive customization for each study and high costs for maintenance, resulting in ideal pre-processed data. Additionally, the authors emphasize that some studies have unique requirements that may not be met by existing commercial solutions. This is supported by numerous papers that describe researchers' experiences in developing their own digital solutions.

A. Bizzego and colleagues describe the development of a custom web-resource called Physiolyze, which was designed to process heart rate variability (HRV) data - one of the most frequently used types of data [Bizzego *et al.*, 2014]. However, processing this type of data requires a high computational capacity, particularly in different life contexts where fast calculations are crucial. The system presented by the researchers is based on two main pillars: pyHRV, an open-source Python toolbox with a range of functions for HRV analysis, and the Galaxy platform - a back-end software tool commonly used for bioinformatics research. The system was successfully tested with different heart rate sensors and demonstrated support for various HRV indexes. The undoubted advantage of the system is that it provides collection of physiological data (heart rate) in the context of real life. But at the same time, Physiolyze does not provide for the construction of research scenarios. And most often it is very important in experimental works.

Another team presented a case study on the development of an educational management information system and provide a detailed analysis of web-service data collection in general [Thukral and Goel, 2012]. The authors emphasize

the positive impact that online data collection methods have brought to the quality of datasets in terms of speed, accuracy, reliability, and completeness. Additionally, the use of web-services has helped to overcome issues related to data transfer in heterogeneous computer environments, time spent on data collection, and data presentation formats. This system is a good example of automation of manual data collection and processing. At the same time, the system assumes data collection only by predetermined blocks and does not assume a branched algorithm. We argue that the information system for assessing sleepiness dynamics should: a) be cyclic (sleepiness should be recorded at equal intervals); b) be branched (cyclicality should be interrupted when a person reports going to bed).

M. Hanbury and colleagues describe how the implementation of a web-based management application improved the data management process for longitudinal studies [Hanbury *et al.*, 2019]. The researchers encountered various challenges, such as geographical remoteness of research sites, variable work hours of the participants, a large number of variables involved in testing, different types of research teams, and a multilingual community of participants. The use of an integrated web-application significantly reduced the amount of time spent on data collection, leaving more time available for the field researchers to communicate with the participants and produce more accurate and detailed reports. The application also facilitated communication between different research teams, enabling them to split up responsibilities and providing real-time updates of gathered information, allowing researchers to adjust the course of the study if necessary. However, access to the Internet was a limitation in some remote parts of the country, and occasional IT support was required from the developers to add users to the system or resolve other programming-related issues. Thus, this case also confirmed that web applications can increase the efficiency of data collection. The most important benefits provided here can be highlighted as: increasing the speed and quality of data collection, getting the possibility of automatic report preparation, and real-time updates on gathered information.

Zhang and colleagues discuss their experience of using digital services in a medical environment [Zhang *et al.*, 2017]. Their goal was to assess the advantages of data collection, project management, and telemonitoring in a hospital setting, and they developed a mobile application called mEDC for electronic data capture. The mEDC consisted of a mobile application and a server-based clinical database that allowed for a range of functions such as creating appointments, signing informed consent, inputting patient data, delivering medicine, and scheduling appointments for biological samples collection. The system was tested with different users from patients to research associates and managers across 14 hospitals, and the researchers received positive feedback, particularly on the real-time data collection and management. The mobile application was particularly effective, as it allowed doctors to input data directly into the system and reduce the time spent on patient records. However, the Internet connection stability was a limitation of the system, and the authors suggest further development to include features such as medical record change tracking and

electronic signature implementation. Consequently, the use of an automated system has saved time while maintaining the information required in this context.

Finally, Canino et al. presented an innovative approach to improving wellness through the integration of geographical data into medical records [Canino *et al.*, 2016]. The authors introduced GeoBlood, a web-based system that collects patients' blood analysis results and their geolocation to investigate the potential correlation between environmental factors and health issues. The system stores clinical information, hospital information, and geographic information, enabling users to access and analyze the collected data. It provides features such as descriptive statistics, data navigation, and data export, allowing users to monitor changes in biological data over time for individual patients, as well as statistical data by age or gender. The system was tested with clinical data and showed promising results, though the authors suggest further improvements are necessary to enhance its functionality. Thus, GeoBlood provides effective access and analysis of the collected data. An important advantage is the possibility both to export data and to obtain descriptive statistics. Also, users have the ability to track the dynamics of the input data, which is important if we consider experimental studies involving large samples.

Thus, the reviewed experience of creating and using automated systems, reflected in scientific publications, shows that such systems are created: (a) either for continuous data recording without regard to the context and the possibility of setting stages; (b) or to automate manual data entry in surveys for different tasks and to automate reporting. The undoubted advantage of using automated systems is the ability to aggregate data and obtain descriptive statistics and reports.

At the same time, the considered systems: (a) do not assume the branched algorithm of the data collection scenario; (b) do not give the possibility of interaction with the system before the stage of authorization. Experimental research in natural conditions with the use of laboratory equipment involves the participant coming to the laboratory to receive instructions and equipment. This stage can also be automated to eliminate the need to contact the participant directly. Authorization is not necessary at this stage, as the potential participant may not 'reach' the stage of the experiment itself. Therefore, we believe that this stage should also be automated and included as part of a comprehensive system for conducting mass psychophysiological research. A branching algorithm is necessary, for example, when studying the dynamics of sleepiness from evening to nighttime because in this case the system needs to disable the cyclic questioning of sleepiness levels when a person reports that he/she is going to bed.

As none of the current automated systems fulfill all the necessary options for our project, we developed our personalized web-application enabling (a) interaction with or without authentication and (b) a branched algorithm for data collection scenarios.

### 3 Web application description

Starting in 2022, the Laboratory of Cyberpsychology conducted a large-scale psychophysiological research study on subjective sleepiness dynamics in adults involving by the date of this paper preparation 230 participants. The study design required each participant to record their electrocardiogram, pulse, and interbit intervals using a Polar H10 sensor from 07:50 PM until 06:00 AM, as well as provide their sociodemographic characteristics, including the presence of chronic conditions, caffeine consumption, the presence of sleep problems, driving experience, sleep and wakefulness habits, and subjective sleep characteristics. In addition, participants were required to fill in three different questionnaires (Karolinska Sleepiness Scale (KSS), Stanford Sleepiness Scale (SSS), and Epworth Sleepiness Scale) being at home. The KSS and SSS are questionnaires that measure situational sleepiness. We used a 7-point scale within the SSS, and a 10-point scale within the KSS. The consistency of these scales was reported by us previously [Demareva *et al.*, 2023].

An automated system called UnnCyberpsy was developed to overcome the challenges of collecting and processing large amounts of data in a recent psychophysiological study on subjective sleepiness dynamics in adults. This web-application was designed using the PHP programming environment and the 'CodeIgniter' microframework (version 4.0), based on the model-view-controller (MVC) pattern. To ensure user-friendliness and compatibility with different devices, the Bootstrap CSS framework (version 5.2) was utilized. The system employs MariaDB for data storage.

UnnCyberpsy provides several key features to facilitate the study. **(1) Firstly**, participants can schedule an appointment to collect the necessary equipment and receive instructions. The system allows them (without authentication) to select a convenient date and time, and sends reminders to their email. The equipment pickup records are based on the availability of Polar H10 sensors, ensuring that participants can only sign up for dates when equipment is available for pickup. Before implementing key feature (1), the experimenters telephoned potential study participants to set a date and time for them to come to the lab. Of the 38 calls, only in 23 cases the date and time could be immediately agreed upon. Seven (7) potential participants did not pick up the phone (probably ignoring calls from unfamiliar numbers), and 8 had already reconsidered participating in the experiment. Thus, the delay that occurs between the application to participate in the experiment and the experimenter's call leads to a loss of motivation to participate in the experiment.

The implementation of key feature (1) made it possible to exclude the participation of the experimenter at the stage of determining the date and time of the respondent's arrival to the laboratory in order to receive instructions and equipment. After the implementation of key feature (1), 90% registration and arrival of potential participants to the laboratory was ensured. Consequently, using UnnCyberpsy in its full version increased the turnout of potential study participants from 60% to 90%.

**(2) Secondly**, the system requires users to create an account that is approved by the administrator, who in this case is the experimenter, before accessing it. Participants pro-

Time	Information to fill out
19:50	• Personal information
19:55	• ESS
20:00 and each 30 m after	• SSS and KSS
Exact time when going to bed	• Time of going to bed
06:00	• SSS and KSS • Sleep quality • Dreams description

Figure 1. The stages of pre-programmed testing.

vide their email addresses as their login IDs, and the system generates a secure, randomly generated password that is sent to their email addresses. The system is pre-programmed for every stage of the testing process. Participants can complete pre-programmed tests with branched algorithm using their own devices, and the system automatically collects, preprocesses, and stores the data. Since the experiment design involves participants filling out tests at specific times, the system automatically opens the tests at predetermined intervals. The stages of pre-programmed testing with timing are presented in Figure 1.

To ensure data precision and uniformity, the system verifies all input fields before storing them in the database. Most queries offer limited response choices, while open-ended ones undergo scrutiny for accurate symbols. When errors are identified, the field is highlighted alongside an error prompt. This method eradicates human mistakes, standardizes participant responses, and bolsters data quality. Elaborate instructions accompany each field, aiding participants in accurately completing the survey. Post-experiment, all participant data undergoes scrutiny and is archived within the system.

Fifteen (15) participants (7% of the sample) completed their cyclic tests with a delay of up to 10-minutes. However, this data remains suitable for analysis. These incidents occurred early in the data collection phase; subsequently, revisions to the web-application instructions eradicated delayed test completions, resulting in a data loss rate of 0% post-update. With the system automatically verifying form accuracy, all stored data stands ready for analysis. This underscores UnnCyberpsy’s capacity to amass high-quality, precise data.

Furthermore, the system’s reliance on a branched algorithm accommodated the unique attributes of each respondent’s sleep-wake patterns, facilitating the assessment of their sleepiness dynamics until the moment of sleep onset.

**(3) Thirdly**, experimenters possess database access, enabling result viewing and download. Participant identification relies on their email address and initial entry date, enabling researchers to synchronize system-recorded data with sensor-collected information. Additionally, the system allows CSV file exports containing user data for further analysis. Designed on these principles, the database grants real-time access to updated results for the research team. An illustrative depiction of the system’s workflow is presented in Figure 2.

Moreover, it’s important to highlight that the quantity of

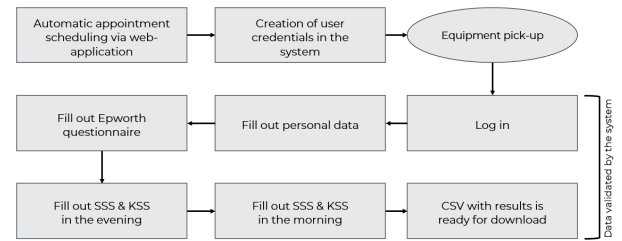


Figure 2. General pipeline of UnnCyberpsy. Squares stand for automatic interactions, and oval stands for person-to-person interaction.

participants capable of partaking in the study within a single day is restricted solely by the availability of cardiac sensor and smartphone kits in the laboratory, currently set at 10. The utilization of this system allows the experimenter to be detached from the actual experiment, thereby accelerating the pace of data acquisition. In previous research on sleepiness (e.g., [Casale *et al.*, 2022]), specialized experimenters were required to oversee adherence to the experiment protocol. Hence, UnnCyberpsy enables swift collection of respondent data concerning sleepiness dynamics and heart rate without any delays. It should be noted that this system can be modified to record physiological or actigraphy data from any devices instead of Polar H10.

It is important to acknowledge that the experimental procedure — characterized by the cyclic completion of tests at specific intervals, the need for independent operation of the equipment, and the absence of an experienced experimenter — could potentially create a stressful environment for participants. Therefore, when developing the UnnCyberpsy system, one of our team’s top priorities was not only to create an automated data collection method to eliminate errors related to human factors but also to make the process as simple and comfortable as possible for participants. This means that the UnnCyberpsy system must act as an experimenter, guiding participants through all stages of the study while remaining intuitive, similar to familiar digital services. In designing the system, we aimed to follow general principles of user experience design, using well-known and established design elements and techniques.

Jakob Nielsen, founder of Nielsen Norman Group — a consulting company specializing in UX — has notably focused on generalizing design elements based on user expectations. Nielsen argues that many common website elements should function according to predictable patterns. Typically, this applies to simple elements like menu icons that display page contents, search bars, or ‘breadcrumbs.’ Familiar scenarios enhance user confidence in completing tasks and, consequently, increase satisfaction with the site. Standardizing core elements allows users to recognize them in the interface, understand how to use them effectively, and avoid missing important but non-obvious features, thereby improving task resolution. Moreover, Nielsen advocated for actively involving users in the design process. He proposed a set of guidelines known as ‘Nielsen’s 10 Heuristics’ — principles for evaluating interface usability from the end user’s perspective. These heuristics include [Nielsen, 1999]:

- Feedback: The user should understand what is happen-

ing with the system, which should provide this information.

- ‘In the user’s world’: The system should communicate in a language the user understands.
- Control and freedom: There should be an option to exit any system state, especially undesirable ones.
- Consistency and standards: The same words and symbols should always have the same meaning or function.
- Error prevention: The system should minimize the conditions under which errors can occur.
- Visibility and clarity: The system should provide all necessary information without requiring the user to recall or search for it.
- Flexibility and efficiency: The system should adapt to both new and experienced users, providing only relevant information.
- Aesthetic and minimalist design: Texts and interface design should be clean and relevant.
- Error handling: Error messages should be clear, precise, and offer solutions.
- Clear documentation: All reference information should be available and written in an understandable language.

In our case, from the start, the system employs familiar registration forms and login credentials. After authorization, users receive brief information about the experiment, including a detailed description of the most complex stages—cyclic test completion—to ensure participants are aware of this aspect in advance. A list of experimental stages is also provided, allowing participants to review the information they will need to enter into the system. Information fields remain inactive until required according to the experiment’s schedule, with the system monitoring when data entry is needed.

The correctness of all data entered is automatically verified. If data is incorrect or fields are missed, the system indicates the specific error location and suggests corrections, marking errors in red. Successful data saving is confirmed by the system with a green message, making the notifications intuitive and identifiable. The cyclic completion of questionnaires is specifically programmed to minimize user error, with the system locking data entry until a set time, displaying the exact time until the form becomes active again. Once this period expires, the field becomes available for entry. All fields are designed to match the page’s ergonomics, with single-choice options presented as bullet points, multiple-choice options marked with checkboxes, predefined parameters available in dropdown lists, and date and time adjustable via arrows or manual entry.

## 4 Case study

The hypothesis of the experiment tested by the system UnnCyberpsy was as follows: the level of subjective sleepiness would increase from evening to nighttime, with its dynamics associated with various characteristics of the participants. The total sample size of the study comprised 225 individuals. Within this article, the results pertaining to 156 participants are presented. These individuals completed the SSS and KSS questionnaires between 20:00 and 22:00 and went

to bed after 22:15. This selection aimed to analyze evening sleepiness dynamics across at least five time points (20:00, 20:30, 21:00, 21:30, and 22:00) and compare it with subjective sleepiness in the morning (at 6:00). The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Faculty of Social Sciences of Lobachevsky State University (Protocol No. 4 from 21 November 2022). Prior to the experiment, each participant read and signed an informed consent.

Student’s t-test for dependent samples was used to assess differences in SSS and KSS scores across various time periods. Pearson’s correlation criterion was applied to evaluate the relationship between SSS and KSS scores at different time points. Additionally, a multifactorial repeated measures analysis of variance was conducted to assess the influence of various factors on KSS and SSS scores.

During the analysis, it was found that 100% of participants completed all experiment stages on time. Indeed, they rated their sleepiness every 30 minutes without any omissions or delays. All participants also indicated the time they went to bed.

The analysis of the time’s impact on subjective sleepiness revealed that time significantly influenced both KSS ( $F = 86.5, p < 0.001$ ) and SSS scores ( $F = 95.9, p < 0.001$ ). Sleepiness increased from 20:00 to 22:00 and remained elevated at 6:00, indicating a distinct trend in SSS and KSS scores. These findings align with data from other studies that similarly confirmed increased sleepiness from evening to nighttime [Smith *et al.*, 2009][Abrahamsen *et al.*, 2022]. Therefore, the data collected using the automated system aligns with data collected through other means. The results obtained confirm the hypothesis that subjective sleepiness levels would rise from evening to nighttime. In our study, scores on both sleepiness assessment scales were closely related at each time point ( $p < 0.001$ ), indicating consistency between SSS and KSS. Thus, scores from these two valid sleepiness scales did not differ when using the automated system.

The combined influence of age group (up to 25 years, 25–35 years, over 35 years) and time on KSS scores ( $F = 3.71, p < 0.01$ ), and the influence of gender on SSS scores ( $F = 5.5, p < 0.05$ ) were identified. BMI (underweight, normal weight, overweight/obese), gender, and daily coffee consumption (0, 1, more than two cups) had an overall impact on KSS scores ( $F = 2.45, p < 0.05$ ). However, BMI did not exhibit independent influence either considering the time factor or for each period separately. The independent effect of age on subjective sleepiness was noted at the 20:00 stage for both SSS ( $F = 5.07, p < 0.01$ ) and KSS ( $F = 3.67, p < 0.05$ ), indicating potential variations in “sleep-wake” patterns among volunteers of different age groups. The independent influence of daily coffee intake on subjective sleepiness was detected only for SSS at the time points 21:30 ( $F = 3.13, p < 0.05$ ) and 22:00 ( $F = 3.48, p < 0.05$ ): individuals who abstained from coffee exhibited lower sleepiness. Interestingly, the number of coffee cups consumed during the experiment day was not associated with subjective sleepiness.

Remarkably, the influence of coffee on subjective sleepiness was rather inconsistent in our study. While the number of coffee cups consumed per day during the experiment did not impact respondents’ subjective sleepiness, daily coffee

intake affected sleepiness late in the evening. This aligns with the varied effects of caffeine. It was suggested that caffeine reduces subjective sleepiness until the body becomes accustomed [Reichert *et al.*, 2022]. Subjective sleep quality was poorer in individuals who regularly consumed more than 8 cups of coffee a day [Sanchez-Ortuno *et al.*, 2005], as caffeine, by blocking the adenosine neuromodulator and receptor system, alters the sleep-wake regulation system and worsens nighttime sleep, with the magnitude of this influence dependent on individual characteristics [Clark and Landolt, 2017].

Consequently, the dynamics of subjective sleepiness were nonlinearly associated with individual characteristics such as BMI, gender, age, and coffee consumption. This affirmed the hypothesis that the dynamics of subjective sleepiness are linked to the characteristics of the individual participants.

## 5 Conclusion

The traditional method of gathering and managing data in a psychophysiological study examining subjective sleepiness dynamics can become daunting, particularly with a substantial participant count and intricate experiment structures. To tackle this hurdle, we devised a personalized web application named UnnCyberpsy, streamlining data collection, storage, and preliminary processing. This innovation allowed our research team to entirely eliminate the risk of data loss, enhance potential participant engagement by 30 percents, and accelerate data gathering, now solely constrained by participant numbers and equipment availability.

Our experience underscores the significant simplification that web applications offer in data collection, thereby streamlining research processes for both investigators and participants. Our forthcoming objective involves expanding UnnCyberpsy's capabilities to include automatic retrieval of heart rate data from Polar H10 sensors via our Android application, 'CyPsy.' The further development of the UnnCyberpsy platform will include the addition of an online informed consent form.

## Declarations

### Acknowledgements

This paper is an extended and revised version of Demareva *et al.* [2024] (forthcoming).

### Authors' Contributions

VD contributed to the conception of this study. VV and IZ performed the experiments and wrote original draft. AD and NN performed the experiments. VD is the main contributor and writer of this manuscript. All authors read and approved the final manuscript.

### Competing interests

The authors declare no conflicts of interests.

## Availability of data and materials

The datasets analysed during the current study are available on request.

## References

- Abrahamsen, A., Weihe, P., Debes, F., and van Leeuwen, W. M. (2022). Sleep, sleepiness, and fatigue on board faroese fishing vessels. *Nature and science of sleep*, pages 347–362. DOI: <https://doi.org/10.2147/NSS.S342410>.
- Alavi, C. B. and Massman, J. D. (2016). Selecting an electronic data capture system. *Urology practice*, 3(3):236–241. DOI: <https://doi.org/10.1016/j.urpr.2015.06.007>.
- Bizzego, A., Mina, M., Zarbo, C., Esposito, G., and Furlanello, C. (2014). Physiolyze: a galaxy-based web service for heart rate variability analysis with online processing. In *2014 8th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, pages 97–98. IEEE. DOI: <https://doi.org/10.1109/ESGCO.2014.6847537>.
- Canino, G., Scarpino, M., Cristiano, F., Mirarchi, D., Tradigo, G., Guzzi, P. H., Cuda, G., and Veltri, P. (2016). Geoblood: a web based tool for geo-analysis of biological data. *Procedia Computer Science*, 98:473–478. DOI: <https://doi.org/10.1016/j.procs.2016.09.077>.
- Casale, C. E., Yamazaki, E. M., Brieva, T. E., Antler, C. A., and Goel, N. (2022). Raw scores on subjective sleepiness, fatigue, and vigor metrics consistently define resilience and vulnerability to sleep loss. *Sleep*, 45(1):zsab228. DOI: <https://doi.org/10.1093/sleep/zsab228>.
- Clark, I. and Landolt, H. P. (2017). Coffee, caffeine, and sleep: A systematic review of epidemiological studies and randomized controlled trials. *Sleep medicine reviews*, 31:70–78. DOI: <https://doi.org/10.1016/j.smrv.2016.01.006>.
- Demareva, V., Viakhireva, V., Demarev, A., and Nazarov, N. (2024). 'unnCyberpsy': A web-application for psychophysiological research of subjective sleepiness dynamics. In *Digital Geography: Proceedings of the International Conference on Internet and Modern Society (IMS 2023)*. Springer.
- Demareva, V., Viakhireva, V., Zayceva, I., Isakova, I., Okhrimchuk, Y., Zueva, K., Demarev, A., Nazarov, N., and Edeleva, J. (2023). Temporal dynamics of subjective sleepiness: A convergence analysis of two scales. *Biological Rhythm Research*, 54(4):369–384. DOI: <https://doi.org/10.1080/09291016.2023.2193791>.
- Hanbury, M. M., Sadeghi, B., Tsergounis, I. E., Gomez-Camacho, R., Manzo, R. D., Rangel, M. I., Alexandrescu, B., and de la Torre, A. (2019). A web-based application to improve data collection in an interventional study targeting childhood obesity: pre-post analysis. *Journal of Medical Internet Research*, 21(1):e10861. DOI: <https://doi.org/10.2196/10861>.
- Jabbar, R., Al-Khalifa, K., Kharbeche, M., Alhajyaseen, W., Jafari, M., and Jiang, S. (2018). Real-time driver drowsiness detection for android application using deep neural networks techniques.



- Procedia computer science*, 130:400–407. DOI: <https://doi.org/10.1016/j.procs.2018.04.060>.
- Mohana, B. and Rani, C. S. (2019). Drowsiness detection based on eye closure and yawning detection. *International Journal of Recent Technology and Engineering*, 8(4):8941–8944. DOI: <http://doi.org/10.35940/ijrte.D9716.118419>.
- Nielsen, J. (1999). User interface directions for the web. *Communications of the ACM*, 42:65–72. DOI: <https://doi.org/10.1145/291469.291470>.
- Phan, A.-C., Trieu, T.-N., and Phan, T.-C. (2023). Driver drowsiness detection and smart alerting using deep learning and iot. *Internet of Things*, 22:100705. DOI: <https://doi.org/10.1016/j.iot.2023.100705>.
- Reichert, C. F., Deboer, T., and Landolt, H.-P. (2022). Adenosine, caffeine, and sleep–wake regulation: state of the science and perspectives. *Journal of sleep research*, 31(4):e13597. DOI: <https://doi.org/10.1111/jsr.13597>.
- Sanchez-Ortuno, M., Moore, N., Taillard, J., Valtat, C., Leger, D., Bioulac, B., and Philip, P. (2005). Sleep duration and caffeine consumption in a french middle-aged working population. *Sleep medicine*, 6(3):247–251. DOI: <https://doi.org/10.1016/j.sleep.2004.10.005>.
- Smith, S. S., Horswill, M. S., Chambers, B., and Wetton, M. (2009). Hazard perception in novice and experienced drivers: The effects of sleepiness. *Accident Analysis & Prevention*, 41(4):729–733. DOI: <https://doi.org/10.1016/j.aap.2009.03.016>.
- Thukral, R. and Goel, A. (2012). Web service interface for data collection. *International Journal of Computer Science*, 9(3):525–530.
- Wong, J. Y. and Lau, P. Y. (2019). Real-time driver alert system using raspberry pi. *ECTI Transactions on Electrical Engineering, Electronics, and Communications*, 17(2):193–203. DOI: <https://doi.org/10.37936/ecti-ec.2019172.215488>.
- Zhang, J., Sun, L., Liu, Y., Wang, H., Sun, N., and Zhang, P. (2017). Mobile device–based electronic data capture system used in a clinical randomized controlled trial: Advantages and challenges. *Journal of Medical Internet Research*, 19(3):e66. DOI: <https://doi.org/10.2196/jmir.6978>.