Assessing Cognitive Load Using Blink Rates and Performance Metrics in Digital Span Tasks

Abhinanda Mannazhath Nandan, Abdalrhman Abdullah, Ahmad Kanawati, Aditya Vikram, Apoorva Gupta, Chinagorom Esther Nomeh, Hend Barakat, Pankhuri

INTRODUCTION

Crafting effective systems involves creating visuals and interactions that do not overwhelm users mentally, allowing them to concentrate on their tasks and utilize their natural abilities effectively. Recent technological advancements have sparked significant interest in understanding various aspects of human cognition, including memory, attention, and reaction, emphasizing measuring cognitive load - the mental effort required for task completion. This knowledge holds vast potential, from enhancing user interfaces to optimizing educational methods and marketing strategies. Monitoring cognitive load in human-computer interaction is essential for tailoring interfaces to users' abilities, ultimately improving their experience and reducing mental fatigue.

Moreover, recent studies have delved into utilizing gaze and pupil metrics to assess cognitive workload across various tasks objectively. These studies have emphasized the importance of task-specific considerations and the complementary role of gaze and pupil metrics in workload assessment, aiming to inform the development of more effective systems and enhance user experiences [1]. Additionally, advancements in eye-tracking technology have enabled researchers to explore natural blinking patterns and their implications for reliable data collection. The variability in blink detection accuracy underscores the critical role of selecting the appropriate technology for dependable results, with recommendations to refine methodologies to ensure consistency and reliability in data collection processes [5].
Furthermore, the exploration of pupil size as an indicator of cognitive effort during tasks has revealed significant insights into the impact of gaze position on measured pupil size. This understanding has led to the proposal of corrective functions to address gaze-position-dependent variations, with implications for improving the accuracy of cognitive pupillometry experiments[4]. Meanwhile, studies investigating the effects of automated driving scenarios on drivers' mental workload have highlighted the potential of pupil diameter as a real-time indicator, suggesting its usefulness in real-world contexts for monitoring cognitive workload and enhancing driving experiences [2].

These collective efforts underscore the interdisciplinary nature of research in understanding human cognition and its applications in various domains, ranging from human-computer interaction with driving automation. By leveraging technological advancements and refining experimental methodologies, researchers aim to advance our understanding of cognitive processes and improve the design of systems to better cater to human capabilities and needs[3].

The selection of these studies is based on the scope of our review and investigation on determination of cognitive load via eye-based parameters: gaze, blinks and pupil diameter. In this research, we are aiming to understand the cognitive load of users in an ideal living environment in their short-term memory using the webcam of their personal device and to establish the relationship of it to the eyelid distance of the users.

METHODS

Equipment and Software

The experiment utilized a specialized software application capable of running both a digital span test and a blink detection module. The software was installed on a standard computing device with an integrated camera. The blink detection functionality relied on Python scripts executed through the Spyder Integrated Development Environment (IDE).

1. Initialization

Participants initiated the experiment by launching the designated software application. This application exclusively accessed the digital span test files, which could only be opened within this program.

2. Digital Span Test

● The software presented a sequence of digits on the screen for a brief period.
● Participants were instructed to memorize the sequence of digits displayed.
● Upon completion of the display, participants were prompted to enter the memorized digits using the keyboard, specifically pressing the spacebar to submit their response.

3. Blink Detection

● Concurrently, the Python script was executed to activate the device's camera.
● The script monitored and recorded the participant's blink activity during the digit memorization task, detecting and counting the number of blinks.

4. Order of Conditions

The experiment comprised three conditions: Baseline, Task, and Recovery.
● Baseline Condition: Participants rested for 5 minutes while their blink rate was recorded to establish a baseline.
● Task Condition: Participants performed the digital span test while their blinks were monitored and recorded.
● Recovery Condition: After completing the digital span test, participants rested again for 5 minutes with continued blink rate monitoring to observe recovery patterns.

5. Repetition and Data Collection

● The entire process was repeated for a total of 15 trials per participant to ensure data reliability.
● Both the results of the digital span tests and the blink counts were
automatically recorded after each trial.

6. Data Exportation

At the conclusion of the session, the software automatically compiled the results into a CSV file. This file included data on the memorization accuracy of the digit sequences and the blink frequency for each trial.

Participants

The study sample will consist of 10 participants. The age of the participants will be between 20 and 30 years old with a mean age of 25 years (SD = 3). The gender distribution will be 50% males and 50% females.

Data Processing

The raw data collected from the experiment will undergo cleaning to ensure its accuracy and reliability. According to Leys et al. (2013), this process involves detecting and removing outliers that deviate more than 3 standard deviations from the mean. This step is essential to reduce the impact of extreme values that could skew the results. Eliminating these outliers will help the data more accurately represent participants' typical performance, thereby improving the validity and reliability of the analysis. Each participant's accuracy was calculated by comparing the total amount of numbers displayed with the correct numbers they identified. Omissions were treated as errors, thus affecting the overall accuracy scores.

RESULTS

Participant	Condition (Digits)	Accuracy (%)
P1	Easy (3 digits)	73.33
P1	Medium (5 digits)	60
P1	Hard (8 digits)	61.54

P2	Easy (3 digits)	86.67
P2	Medium (5 digits)	60
P2	Hard (8 digits)	62.96
P3	Easy (3 digits)	0
P3	Medium (5 digits)	0
P3	Hard (8 digits)	0
P4	Easy (3 digits)	83.33
P4	Medium (5 digits)	50
P4	Hard (8 digits)	20.51
P5	Easy (3 digits)	0
P5	Medium (5 digits)	0
P5	Hard (8 digits)	0
P6	Easy (3 digits)	86.67
P6	Medium (5 digits)	84.44
P6	Hard (8 digits)	98.22

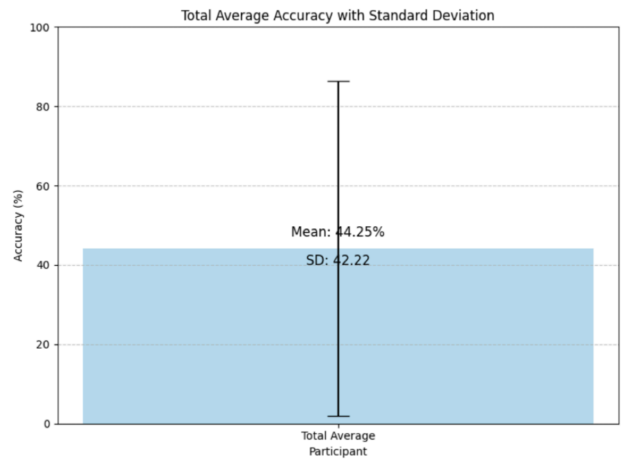
Overall Performance Metrics

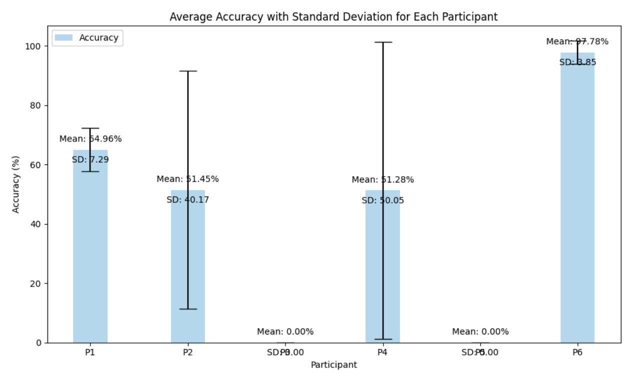
Here, we computed the mean and standard deviation of accuracies across all participants for each condition. The results are as follows:

- Easy Condition: Average accuracy was approximately 58.89% with a high standard deviation of 46.46%, indicating significant variability in participant performance.
- Medium Condition: Average accuracy decreased to around 45.64% with a standard deviation of 38.99%, reflecting increased task difficulty.
- Hard Condition: Average accuracy further dropped to 28.21%, with a standard deviation of 42.60%, showcasing substantial challenges faced by participants in the hardest setting.

Visualization

A bar chart was created to illustrate the average accuracies for each condition, with error bars representing the standard deviations.





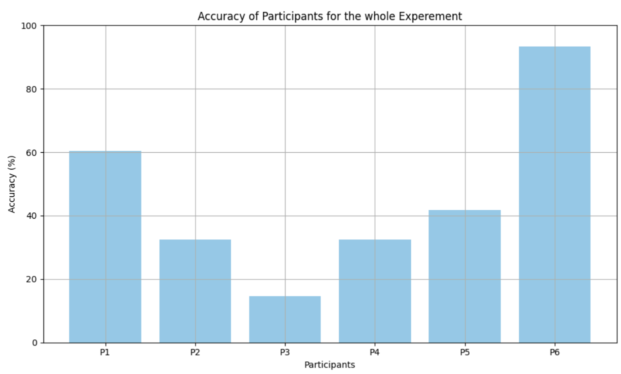
Overall Accuracy of Each Participant

The overall accuracy for each participant was calculated across all conditions. The results are as follows:

Participant	Overall Accuracy(%)
P1	64.96
P2	51.45
P3	0
P4	51.28
P5	0
P6	97.78

Observation

The total average accuracy of 44.25% across all participants and digit sequence lengths (3, 5, and 8 digits) highlights the overall performance of the transcription or recognition process. This metric indicates the general level of correctness achieved. The relatively modest average suggests that there is substantial room for improvement in the data processing methodologies, potentially leading to more accurate and reliable outcomes.



Count the blinks of each participant in each condition

We'll start by examining the provided blink rate data for each participant under the three conditions: easy, medium, and hard.

Participant	Blink rate	condition
p1	57	easy
p1	l264	easy
p1	75	medium
p1	212	medium
p1	43	hard
p1	211	hard
p2	40	easy
p2	596	easy
p2	552	medium
p2	41	hard
p2	574	hard
p3	32	easy
p3	508	easy
p3	517	medium
p3	517	hard
p4	678	easy
p4	117	easy
p4	409	medium
p4	38	hard
p4	547	hard
p5	580	easy
p5	88	easy
p5	366	medium
p5	62	hard
p5	682	hard
p6	17	easy
p6	858	easy
p6	725	medium
p6	329	hard

Overall Performance Metrics for Blink Rates

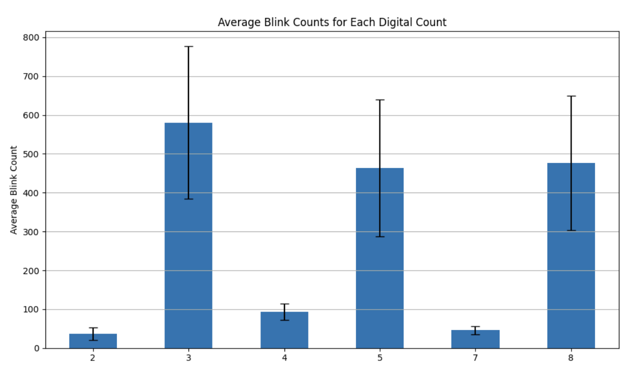
Data Summary: The dataset represents the average blink rates for each participant across three conditions: easy, medium, and hard. These averages were calculated based on the number of blinks recorded during tasks of varying digit counts, which were classified into the three conditions based on their complexity.

Participant	Condition	Average blink rate
p1	easy	160.5
p1	medium	143.5
p1	hard	127
p2	easy	318
p2	medium	552
p2	hard	307.5

p3	easy	270
p3	medium	517
p3	hard	517
p4	easy	397.5
p4	medium	409
p4	hard	292.5
p5	easy	334
p5	medium	366
p5	hard	372
p6	easy	437.5
p6	medium	725
p6	hard	329

Observations

Across all participants, we observe an increase in blink rates as the tasks increase in complexity from easy to medium but a varied response from medium to hard. This inconsistency might suggest that while the medium tasks require more cognitive effort compared to easy tasks, the hardest tasks might involve a level where the cognitive load stabilizes or involve different strategies that do not linearly increase blink rates.



DISCUSSION

The primary objective of this study was to assess the cognitive load of users in an ideal living environment using the webcam of their personal device. This was done by analyzing the participants' performance on a digital span test and monitoring their blink rates under varying conditions of task difficulty. The results indicate a clear relationship between task difficulty, cognitive load, and both accuracy and blink rates.

Performance Accuracy

Participants exhibited significant variability in accuracy across the three conditions (easy, medium, and hard). The highest average accuracy was observed in the easy condition (58.89%), followed by the medium condition (45.64%), and the lowest in the hard condition (28.21%). The decreasing accuracy with increasing task difficulty aligns with existing literature on cognitive load, suggesting that more complex tasks demand greater mental effort, leading to a higher likelihood of errors.

Blink Rates

The analysis of blink rates provided additional insights into cognitive load. For most participants, blink rates increased from the easy to medium conditions, indicating heightened cognitive effort. However, the response from medium to hard conditions was varied, suggesting that cognitive load might involve different coping strategies as task difficulty reaches its peak.

Understanding Cognitive Load

The findings from this study underscore the importance of understanding cognitive load in designing user interfaces and systems. By recognizing that increased task complexity leads to higher cognitive load and reduced performance accuracy, designers can create more user-friendly interfaces that minimize unnecessary cognitive strain. This is particularly relevant in high-stakes environments such as driving automation and educational platforms, where cognitive overload can have critical consequences.

Blink Rate as an Indicator

The use of blink rate as an indicator of cognitive load offers a non-invasive and real-time method for monitoring mental effort. This approach could be particularly useful in real-world applications where continuous assessment of cognitive load is necessary. For instance, in automated driving scenarios, monitoring drivers' blink rates could provide valuable information about their mental state and readiness to take control if needed.

Methodological Considerations

The significant variability observed in the blink rates across conditions highlights the need for refined methodologies in data collection and analysis. Future research should focus on standardizing blink detection techniques to ensure consistency and reliability. Additionally, considering individual differences in blinking patterns and cognitive processing is crucial for developing more accurate predictive models.

Limitations

This study has several limitations that should be addressed in future research. The sample size of 10 participants is relatively small, limiting the generalizability of the findings. Moreover, the age range (20-30 years) and equal gender distribution, while controlled, may not represent broader populations. Future studies should include larger, more diverse samples to enhance the robustness of the results.

Additionally, the experimental setup using a standard computing device with an integrated camera may not capture the full range of natural blinking behaviors compared to more advanced eye-tracking technologies. Refining the blink detection software and incorporating more sophisticated equipment could improve data accuracy.

In Summary, This study demonstrates the potential of using blink rate as an indicator of cognitive load in digital environments. The clear relationship between task difficulty, cognitive load, and performance accuracy provides a foundation for further exploration and refinement of cognitive load assessment methods. By leveraging these insights, designers can create more effective, user-friendly systems that align with human cognitive capabilities, ultimately enhancing user experience and reducing mental fatigue.

REFERENCES

[1] Gorin, H., Patel, J., Qiu, Q., Merians, A., Adamovich, S., & Fluet, G. (2024). A Review of the Use of Gaze and Pupil Metrics to Assess Mental Workload in Gamified and Simulated Sensorimotor Tasks. Sensors, 24(6), 1759.

[2] Radhakrishnan, V., Louw, T., Gonçalves, R. C., Torrao, G., Lenné, M. G., & Merat, N. (2023). Using pupillometry and gaze-based metrics for understanding drivers’ mental workload during automated driving. Transportation research part F: traffic psychology and behaviour, 94, 254-267.

[3] Mathôt, S., & Vilotijević, A. (2023). Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. Behavior Research Methods, 55(6), 3055-3077.

[4] Gagl, B., Hawelka, S., & Hutzler, F. (2011). Systematic influence of gaze position on pupil size measurement: analysis and correction. Behavior research methods, 43, 1171-1181.

[5] Culemann, W., Neuber, L., & Heine, A. (2023). Pupil vs. Eyelid: Evaluating the Accuracy of Blink Detection in Pupil-based Eye Tracking Devices. Procedia Computer Science, 225, 2008-2017.