Evaluating a Gaze-Controlled Navigation Interface: A User-Centered Approach

Parastoo Azizinezhad[1], Anirban Chowdhury[1]

***Abstract*— The design and acceptability of new products are essential factors in their development, particularly for assistive devices that users rely on for everyday tasks. Gathering user feedback is vital to ensuring that these devices are intuitive, simple to use, and not overwhelming. This study involved 19 participants who tested a newly developed gaze-controlled interface, navigated through predefined paths, and completed questionnaires. This study involved two predefined paths, with participants navigating the first path twice before proceeding to the second path. On average, participants reported low mental, physical, and temporal demands. While younger participants finished the tasks more quickly, participants over the age of 50 showed notable improvement in performance during the second round, likely due to greater familiarity with the system, suggesting that the interface is easy to learn. Furthermore, a significant correlation was found between lap completion time and self-reported mental demand (r = 0.609, p = 0.007), indicating a connection between task performance and perceived cognitive load. These results emphasize the importance of designing assistive technology that is accessible and effective for users of varying ages and abilities.**

***Index Terms*— Gaze-Control, Mental Demand, Assistive Technology**

I. Introduction

While modern assistive technologies (ATs) have significantly enhanced the quality of life for individuals with disabilities, the challenge of AT abandonment remains a major obstacle to widespread and long-term adoption [1]. The World Health Organization projects that by 2030, nearly 2 billion people will require assistive devices [2]. This report highlights the urgent need to reduce abandonment rates in future AT developments. Traditionally, abandonment has been attributed to factors such as physical comfort, task efficiency, and ease of use. However, with the integration of artificial intelligence and robotics, new challenges emerge—particularly in ensuring that these technologies can effectively manage social interactions and adapt to users' fluctuating cognitive load and fatigue levels [3].

Existing ATs often struggle to adapt to the ever-changing conditions of human users, which reduces their overall effectiveness. To enhance usability and long-term acceptance, future ATs must be designed to dynamically adjust their behavior in real time which ensures a seamless and intuitive user experience. Developing a human-machine interaction framework that prioritizes adaptability and responsiveness to individual needs can significantly improve the suitability of ATs, ultimately leading to broader adoption and increased user satisfaction.

Research indicates that pupil dilation increases during effortful decision-making which reflects heightened cognitive load, which can significantly impact user performance [4]–[8]. Moreover, pupil size's rapid responsiveness to neural activity highlights its potential as a valuable physiological measure for dynamically adapting assistive systems to users' cognitive states and needs. Reducing cognitive load is crucial in system design, particularly for assistive devices, as users often rely on them for extended periods while managing health-related challenges [9]. Despite advancements in human-computer interaction, there is still a significant gap in developing adaptive systems that can dynamically adjust to users' cognitive workload. Gaze-controlled assistive devices, utilizing eye-tracking technology, have been widely used to support individuals with severe disabilities [10]–[12]. Eye-tracking and pupillary responses have garnered attention across fields such as education [13], psychology [14], [15], marketing [16], and communication [17].

Self-reported measures are crucial in validating users' perceptions, emotions, and overall evaluation of assistive technologies, offering valuable insights into their experiences and interactions with these systems. One widely used tool for assessing task demands is the NASA Task Load Index (NASA TLX) [18], [19], a standardized questionnaire employed in various studies to evaluate cognitive and physical challenges [20]. The NASA TLX evaluates six key factors: mental demand, physical demand, temporal demand, performance, effort, and frustration, providing a comprehensive understanding of the user experience during task completion. Another widely used questionnaire in assistive technology research is the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) [21], which assesses user satisfaction across 12 key factors, such as safety, comfort, durability, effectiveness, and ease of use [22]. Together, these tools help provide a well-rounded understanding of the tasks' demands and users' satisfaction with the technology and guide improvements and enhancements in assistive devices.

With the growing popularity of eye-tracking technology, alongside other biosignal-based modalities for human-computer interaction—such as brain-computer interfaces (BCIs) [23], [24]—future research may explore the integration of eye-tracking with BCI in a hybrid-BCI architecture [25]. This fusion could enhance the adaptability and responsiveness of assistive systems, enabling more efficient and intuitive interaction for users with severe motor impairments. Despite extensive research into autonomous assistive technologies, their widespread adoption among target users remains limited [26]. One critical factor in improving user acceptance is recognizing and adapting to dynamic changes in mental states during interactions with these systems. Understanding these fluctuations in cognitive load and emotional states can enhance the responsiveness and overall effectiveness of assistive technologies, leading to more intuitive and user-friendly experiences [27].

Eye-tracking has been extensively used in psychological studies to understand human cognition and areas of interest. Majaranta et al. [28] and Krafka et al. [29] explored its

[1]P. Azizinezhad and A. Chowdhury are with the School of Computer Science and Electronic Engineering, University of Essex, Colchester, United Kingdom. (Email: `p.azizinezhad@essex.ac.uk`; `a.chowdhury@essex.ac.uk`)

TABLE I
STEPS FOR EACH MOVEMENT

Step	Description
Lateral rotation	45° per action using the left and right buttons
Linear movement	10 cm per action using the forward and backward buttons

applications in providing insights into cognitive processes. In the context of mental workload estimation, numerous studies have focused on the relationship between pupil dilation and cognitive load. Palinko et al. [30] and Ahmad et al. [31] examined how pupil responses can serve as an indicator of mental workload. Guo et al. [32] highlighted that time pressure and system latency, alongside task specificity, influence mental workload, making gaze information, including pupil diameter, a critical factor in workload estimation. Additionally, researchers have delved into alternative methods to assess cognitive load, such as analyzing power and frequency at the mean frequency, as shown by [33].

The influence of mental activities on pupil diameter is well-documented. For example, Van Der Wel et al. [34] found that increased mental load is generally associated with pupil dilation, whereas fatigue leads to pupil constriction, a finding supported by [35]. Furthermore, pupil diameter has been used to differentiate cognitive processes, such as decision-making and sustained attention that highlighted its relevance in assessing user engagement and cognitive states [36]. These findings underscore the potential of pupil dilation as a valuable indicator for understanding mental workload and cognitive focus, particularly in environments with fluctuating cognitive demands.

This paper presents the results of a navigation experiment conducted with a novel gaze-controlled interface. Nineteen healthy volunteers participated in the study, navigating through two distinct paths. The subsequent sections provide an overview of the methodology and system design, followed by the experimental protocol. Section III presents the findings, along with conclusions and recommendations for future work.

II. Material and Methods

In this study, we utilized a mobile robot navigated through an eye-tracking-based interface to evaluate its usability and effects on pupil response. Additionally, we assessed self-reported user experience using standardized questionnaires such as NASA-TLX and Quest 2.0. This approach allowed us to analyze both objective physiological changes and subjective workload and satisfaction ratings and provided a comprehensive understanding of the interface's effectiveness.

A. System Design

A live video feed was centrally displayed on the interface. It provides real-time visual feedback of the robot's position and its surrounding environment. This feed served as the primary reference for participants and enabled them to accurately navigate the robot by observing its movements and adjusting their gaze-based commands accordingly. This real-time feedback was crucial for maintaining spatial awareness and ensuring precise control.

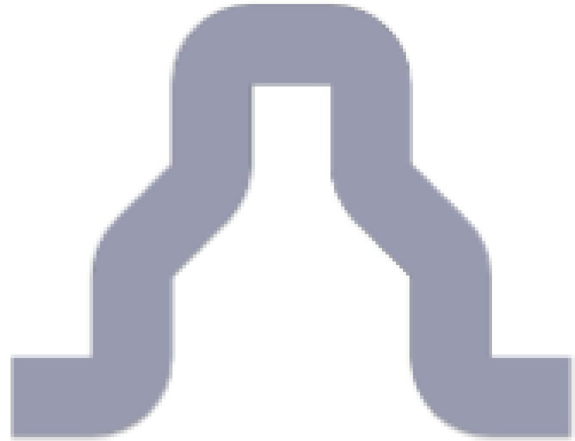

Fig. 1. The image of Path 1, designed for the experiment. All participants navigated this path twice.

The eye-tracking system continuously monitored the participant's gaze and translated it into cursor movements on the screen. This enables hands-free interaction. To issue a command, the participant needed to fixate on a specific button for at least 2 seconds, after which the system would trigger the corresponding action. The "dwell time" rule was implemented for two key reasons:

1) To reduce errors caused by brief, unintentional eye movements, preventing unintended commands.
2) To ensure each command was fully executed—such as completing a 10 cm movement or a 45° turn—before the robot processed a new input, maintaining smooth and controlled navigation.

By implementing this approach, the system enhanced both accuracy and user confidence, making gaze-based robot control more reliable and practical for real-world applications.

B. Experimental Protocol

A total of 19 healthy adults were recruited to participate in the study. All participants provided written informed consent before taking part, and the study was reviewed and approved by the University of Essex Ethics Subcommittee 3 (Ref: ETH2223-2300). To ensure consistency and reliability in the results, participants with uncorrected vision problems were excluded. This exclusion criterion was essential to minimize potential confounding factors, such as variations in visual acuity affecting eye-tracking accuracy or prior familiarity with similar systems influencing performance.

Participants were asked to complete a navigation task that involved two different routes:

- Path 1 (P1): This was the shorter path as shown in 1, which was repeated twice (denoted P1R1 and P1R2) to help participants become more familiar with the interface. The challenge in this path arose during the second half, where participants needed to focus on determining the robot's direction (right or left). This

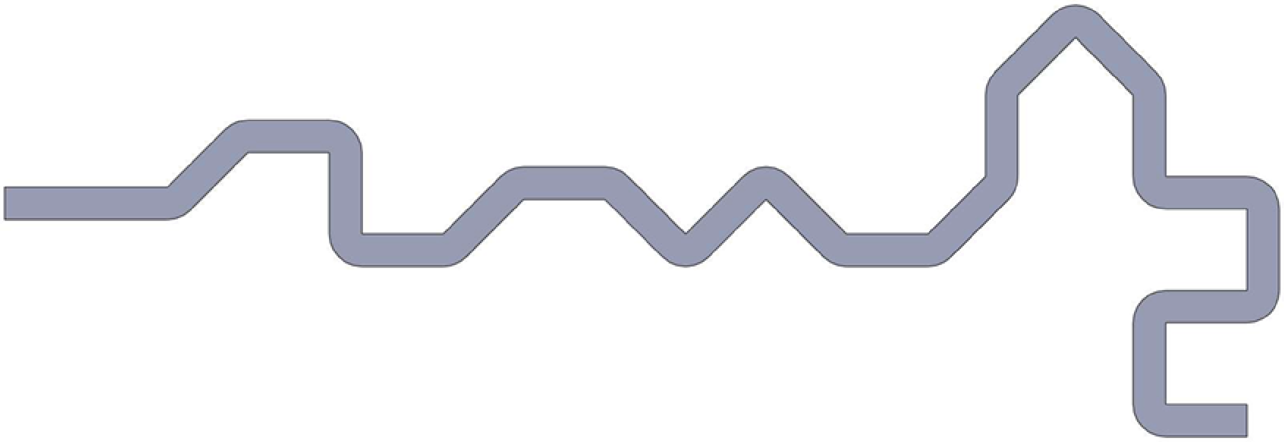

Fig. 2. The image of the longer path (path 2) used for the experiment.

required concentration and careful decision-making, as small errors could disrupt navigation.

- Path 2 (P2): This was the longer route as shown in 1, designed to increase cognitive and physical fatigue by incorporating several sharp turns and varying angles. The intention behind this path was to provide a more demanding task that pushed the participants' endurance and attention over a prolonged period.

Both paths were designed to align with the robot's movement limitations and ensured that the turns adhered to a $45°$ rotation step, which is the maximum turn angle the robot could perform accurately. The order of the paths was fixed (P1 first, followed by P2) to prevent any learning bias that might arise if participants were exposed to the more difficult route (P2) first. These design choices were carefully considered to balance ease of use with progressively increasing task complexity. This ensures participants had a fair opportunity to learn and adjust to the system before encountering more challenging tasks.

Before each trial, the eye tracker was calibrated using a 9-point setup to ensure accurate tracking of the participants' gaze. Following calibration, participants were instructed to guide the robot from a starting point to a designated endpoint and asked to keep the robot on track throughout the navigation task. The task required participants to focus on both the robot's movement and maintaining its correct direction, fostering active engagement. The system recorded performance data in real-time, including metrics such as the time taken to complete the task and the error rate (e.g., deviations from the correct path or incorrect directional inputs). Additionally, gaze tracking data was collected continuously at a rate of 60 Hz which provides detailed information about the participants' gaze patterns during the task. This allowed for an in-depth analysis of how participants interacted with the interface, including how their gaze directed the robot's movements and whether any visual or cognitive challenges arose during the navigation process.

After completing the navigation task, participants were asked to fill out several questionnaires to provide feedback on their experience. These included the NASA TLX, Quest 2.0, and a set of system feedback questions. The NASA TLX was used to assess the mental and physical demands of the tasks, helping to gauge how challenging the participants found the navigation and control process. The Quest 2.0 questionnaire provided insights into the user experience, specifically evaluating aspects like comfort and ease of use of the interface. Participants were categorized into two groups based on age: under 50 years old and over 50 years old. Additionally, the participants were divided into two groups based on their use of vision correction: those who wore glasses or corrective lenses during the experiment and those who did not. This categorization allowed for an analysis of how factors like age and vision correction might influence performance and user experience with the gaze-controlled system. The summary of key details of the experimental setup is shown in Table II.

By collecting both performance and gaze data, the study sought to gain valuable insights into the relationship between user behavior and system efficiency, which are essential for evaluating the usability and effectiveness of the gaze-controlled robot interface. Additionally, by considering various participant groups—categorized by age and vision correction—the study aimed to understand the impact of individual differences on task performance and overall user satisfaction. This approach provided a more nuanced understanding of how diverse populations interact with the system, highlighting potential variations in usability across different user demographics.

III. Experimental results

As shown in Figure 3, familiarity with Path 1 yielded the best results in reducing the completion time in the second round for older participants, with a correlation coefficient of $r = 0.518$ and a statistically significant p-value of 0.028. This suggests that the repetition of the shorter route helped older participants perform the task more efficiently, possibly due to increased comfort and familiarity with the interface, leading to quicker navigation times in the subsequent trial. Such findings highlight the potential benefits of task repetition in improving performance, particularly for users who

TABLE II
SUMMARY OF EXPERIMENTAL SETUP, DEVICES, AND DATA COLLECTION METHODS

Category	Details
Device	**Eye-Tracking System:** Tobii Pro Nano (60 Hz sampling rate) **Robot:** Turtlebot Burger 3
Questionnaires	**NASA TLX:** Measures mental and physical demand of tasks **Quest** 2.0**:** Measures user experience, comfort, and usability of the system **System Feedback:** Questions related to participant feedback on the system
Paths	**Path** 1 **(P1):** Short path, repeated twice (P1R1, P1R2) for familiarization with the interface. Challenge in the second half—requires concentration for right/left direction. **Path** 2 **(P2):** Longer path, designed to increase fatigue with various turn angles and longer distance.
Task Details	**Navigation Task:** Participants were asked to guide the robot from a starting point to a designated endpoint, staying on track. **Data Collected:** Performance data (time taken, error rate) and gaze tracking data (60 Hz rate) during real-time task execution.
Participant Demographics	**Total Participants:** 19 healthy adults **Exclusion Criteria:** Participants with uncorrected vision problems.
Groups for Analysis	**Age Group:** Under 50 years old, Over 50 years old. **Vision Correction Group:** Participants who wore glasses or corrective lenses, and those who did not.
Calibration Process	**Eye Tracker Calibration:** 9-point setup before each trial to ensure accurate gaze tracking.
Actions/Tasks	**Command Selection:** Participants guided the robot by fixing their gaze on specific buttons for 2 seconds to issue commands. **Turn Limitations:** Robot movement aligned with $45°$ rotation step.

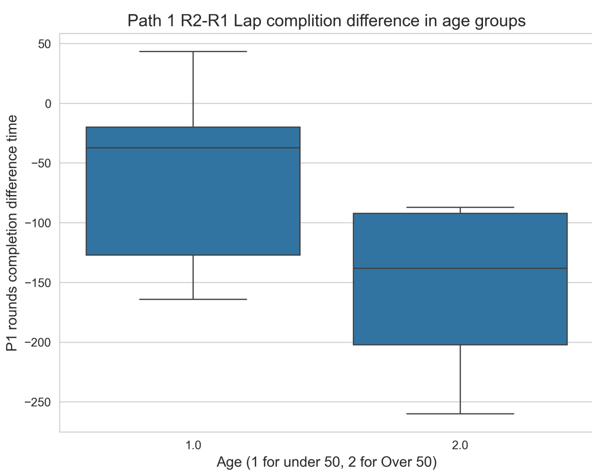

Fig. 3. The box plot presenting the difference of lap completion time of round 2 and round 1 on the path 1 based on the age groups.

may take longer to adapt to novel control methods.

Figure 4 shows a significant correlation between lap completion time and the self-reported mental demand on the NASA TLX questionnaire (Mental Demand vs. Lap Completion Time: $r = 0.609$, $p = 0.007$). The NASA TLX scale ranges from -10 to $+10$, with -10 representing very low demand and +10 indicating very high demand. This suggests that as the mental demand reported by participants increased, so did the time required to complete each lap. The positive correlation highlights the cognitive load involved in navigating the robot and it means that tasks perceived as more mentally demanding may lead to longer task completion times. This finding provides valuable insight into how mental effort affects performance, especially when interacting with a gaze-controlled interface.

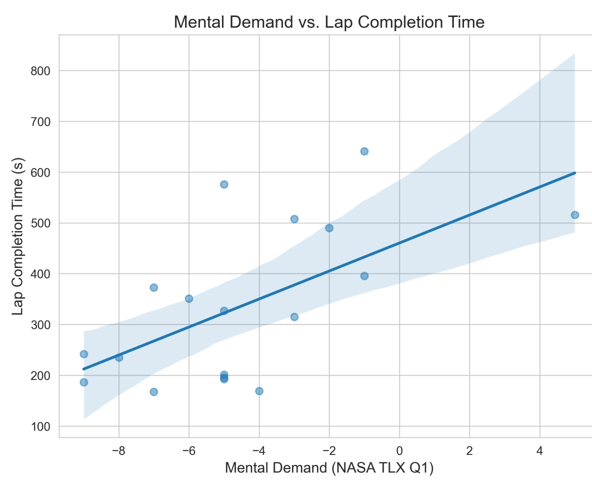

Fig. 4. Lap completion times for Path 1 in Round 1, alongside participants' self-reported mental demand ratings.

On average, participants reported that the system did not induce much demand. The overall results from the NASA TLX questionnaire are shown on Table III.

As mentioned in Table III, these negative values indicate that participants did not perceive the task as particularly demanding in terms of mental and physical effort. The mental demand score of -4.44 suggests that users did not find the task excessively mentally taxing, and the physical demand score of -2.94 further supports that the gaze-controlled system did not require significant physical effort. The temporal demand score of -4.28 indicates that partic-

TABLE III

OVERALL RESULTS FROM NASA TLX QUESTIONNARIE.

Demand	Values (-10 to 10)
Mental Demand	-4.44
Physical Demand	-2.94
Temporal Demand	-4.28
Performance	-5.39
Effort	-1.22
Frustration	-6.89

ipants did not feel rushed or under time pressure during the task. The performance score of -5.39 suggests that, although participants felt the system was somewhat effective, there was still room for improvement in terms of achieving optimal performance. Effort received the lowest score at -1.22, implying that participants felt the system did not require much effort to use, which is a positive indicator for ease of use. However, frustration was the most negative score at -6.89, highlighting that some participants might have experienced frustration, possibly due to challenges in interaction or occasional delays in command execution.

Overall, the results suggest that the system is fairly intuitive and low in demand in terms of cognitive and physical effort, but there may be room for improvement in the performance and frustration areas. It seems that participants valued a system that was easy to use and did not require much mental or physical effort, but there may have been some usability challenges that contributed to frustration. Improving the responsiveness of the system and ensuring that users have more control could enhance the user experience and reduce the frustration score.

Among the 12 items included in the Quest 2.0, participants prioritized 'Comfort,' 'Effectiveness,' and 'Ease of Use' as the most important features when selecting an assistive device. These preferences suggest that users value devices that not only provide optimal performance but also ensure a high level of comfort and simplicity in their operation, which is crucial for long-term use and user satisfaction.

In both age groups, participants showed improvement in performance during the Path 1 on the second round, as indicated by a reduction in lap completion times which is shown in Figure 5. For the under-50 group, the average completion time for Path 1, Round 1 was approximately 230 seconds, which decreased to around 200 seconds in Round 2, which reflects improved efficiency with practice. However, the under-50 group took approximately 500 seconds to complete Path 2, which was a more challenging route due to its length and various turns. For the over-50 group, Path 1, Round 1 took significantly longer, with an average completion time of about 550 seconds. In Round 2, there was a noticeable improvement, as the completion time decreased to approximately 400 seconds. Despite this improvement, the over-50 group still performed considerably slower in Path 2, which is taking around 1150 seconds to complete the more complex route.

Figure 5 illustrates that while both age groups demonstrated learning and adaptation between the first and second

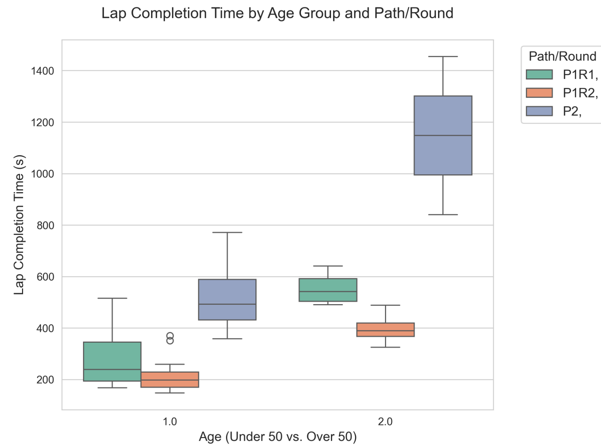

Fig. 5. This plot illustrates lap completion times categorized by age group, where 1 represents participants under 50 years old and 2 represents participants over 50 years old.

rounds for Path 1, older participants faced more difficulty in completing the longer and more intricate Path 2. This may be due to various factors, such as the increased cognitive or physical demands associated with navigating a more complex route, which likely contributed to the longer completion times for the over-50 group. Nonetheless, the improvement between rounds indicates that both groups were able to adjust and become more efficient with the system over time.

IV. Conclusion

This study highlights the importance of user-centered design in the development of assistive technologies, with a particular focus on gaze-controlled interfaces. Our findings show that the proposed system successfully reduced cognitive, physical, and temporal demands on users and made it a promising tool for assistive navigation. Although younger participants completed the tasks more quickly, older participants demonstrated significant improvements in performance during repeated trials ($r = 0.518$, $p = 0.028$) which suggested that the system is intuitive and easy to learn. Additionally, the significant correlation between lap completion time and self-reported mental demand ($r = 0.609$, $p = 0.007$) emphasizes the influence of cognitive load on navigation efficiency.

These results suggest that gaze-controlled interfaces have strong potential as accessible and effective solutions for individuals in need of assistive technology. Future research will focus on using the improved interface to control wheelchair movement, with an emphasis on enhancing its adaptability and long-term usability. Specifically, the integration of real-time adjustments will be explored to further optimize the user experience.

References

[1] H. Petrie, S. Carmien, and A. Lewis, "Assistive technology abandonment: research realities and potentials," in *Computers Helping People with Special Needs: 16th International Conference, ICCHP 2018, Linz, Austria, July 11-13, 2018, Proceedings, Part II 16*. Springer, 2018, pp. 532–540.

[2] A. A. Agency, "Assistive technology in the uk 2022 report," https://analytics.dkv.global/AssistiveTech-in-UK-2022.pdf, 2022, [Accessed: 26.07.2023].

[3] A. Malhotra, "Socially intelligent affective ai," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 11, 2022, pp. 12 888–12 889.

[4] H. Simpson and S. M. Hale, "Pupillary changes during a decision-making task," *Perceptual and Motor Skills*, vol. 29, no. 2, pp. 495–498, 1969.

[5] J. W. De Gee, T. Knapen, and T. H. Donner, "Decision-related pupil dilation reflects upcoming choice and individual bias," *Proceedings of the National Academy of Sciences*, vol. 111, no. 5, pp. E618–E625, 2014.

[6] J. Engström, G. Markkula, T. Victor, and N. Merat, "Effects of cognitive load on driving performance: The cognitive control hypothesis," *Human factors*, vol. 59, no. 5, pp. 734–764, 2017.

[7] E. H. Hess and J. M. Polt, "Pupil size in relation to mental activity during simple problem-solving," *Science*, vol. 143, no. 3611, pp. 1190–1192, 1964.

[8] R. Mitra, K. S. McNeal, and H. D. Bondell, "Pupillary response to complex interdependent tasks: A cognitive-load theory perspective," *Behavior Research Methods*, vol. 49, pp. 1905–1919, 2017.

[9] S. Chen and J. Epps, "Using task-induced pupil diameter and blink rate to infer cognitive load," *Human–Computer Interaction*, vol. 29, no. 4, pp. 390–413, 2014.

[10] M. Borgestig, J. Sandqvist, R. Parsons, T. Falkmer, and H. Hemmingsson, "Eye gaze performance for children with severe physical impairments using gaze-based assistive technology—a longitudinal study," *Assistive technology*, vol. 28, no. 2, pp. 93–102, 2016.

[11] C.-S. Hwang, H.-H. Weng, L.-F. Wang, C.-H. Tsai, and H.-T. Chang, "An eye-tracking assistive device improves the quality of life for als patients and reduces the caregivers' burden," *Journal of motor behavior*, vol. 46, no. 4, pp. 233–238, 2014.

[12] Y. K. Meena *et al.*, "Emohex: An eye tracker based mobility and hand exoskeleton device for assisting disabled people," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2016, pp. 002 122–002 127.

[13] J. L. Rosch and J. J. Vogel-Walcutt, "A review of eye-tracking applications as tools for training," *Cognition, technology & work*, vol. 15, pp. 313–327, 2013.

[14] J. F. Cavanagh, T. V. Wiecki, A. Kochar, and M. J. Frank, "Eye tracking and pupillometry are indicators of dissociable latent decision processes." *Journal of Experimental Psychology: General*, vol. 143, no. 4, p. 1476, 2014.

[15] D. Chatterjee *et al.*, "Real time estimation of task specific self-confidence level based on brain signals," *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 19 203–19 217, 05 2021.

[16] G. van Loon, F. Hermsen, and M. Naber, "Predicting product preferences on retailers' web shops through measurement of gaze and pupil size dynamics," *Journal of Cognition*, vol. 5, no. 1, 2022.

[17] Y. K. Meena *et al.*, "A hindi virtual keyboard interface with multimodal feedback: A case study with a dyslexic child," in *Proceedings of the 32nd Human Computer Interaction Conference*, 07 2018.

[18] S. Hart, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," *Human mental workload/Elsevier*, 1988.

[19] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, no. 9. Sage publications Sage CA: Los Angeles, CA, 2006, pp. 904–908.

[20] S. Said, M. Gozdzik, T. R. Roche, J. Braun, J. Rössler, A. Kaserer, D. R. Spahn, C. B. Nöthiger, and D. W. Tscholl, "Validation of the raw national aeronautics and space administration task load index (nasa-tlx) questionnaire to assess perceived workload in patient monitoring tasks: pooled analysis study using mixed models," *Journal of medical Internet research*, vol. 22, no. 9, p. e19472, 2020.

[21] L. Demers, R. Weiss-Lambrou, and B. Ska, "The quebec user evaluation of satisfaction with assistive technology (quest 2.0): an overview and recent progress," *Technology and disability*, vol. 14, no. 3, pp. 101–105, 2002.

[22] L. Demers, M. Monette, Y. Lapierre, D. Arnold, and C. Wolfson, "Reliability, validity, and applicability of the quebec user evaluation of satisfaction with assistive technology (quest 2.0) for adults with multiple sclerosis," *Disability and rehabilitation*, vol. 24, no. 1-3, pp. 21–30, 2002.

[23] A. Chowdhury and J. Andreu-Perez, "Clinical brain–computer interface challenge 2020 (cbcic at wcci2020): Overview, methods and results," *IEEE Transactions on Medical Robotics and Bionics*, vol. 3, no. 3, pp. 661–670, 2021.

[24] P. Saideepthi *et al.*, "Sliding window along with eegnet-based prediction of eeg motor imagery," *IEEE Sensors Journal*, vol. 23, no. 15, pp. 17 703–17 713, 2023.

[25] A. Chowdhury, A. Dutta, and G. Prasad, "Corticomuscular co-activation based hybrid brain-computer interface for motor recovery monitoring," *IEEE Access*, vol. 8, pp. 174 542–174 557, 2020.

[26] P. Saideepthi, Chowdhury *et al.*, "Sliding window along with eegnet-based prediction of eeg motor imagery," *IEEE Sensors Journal*, vol. 23, no. 15, pp. 17 703–17 713, 2023.

[27] M. F. Hinss, A. M. Brock, and R. N. Roy, "Cognitive effects of prolonged continuous human-machine interaction: The case for mental state-based adaptive interfaces," *Frontiers in Neuroergonomics*, vol. 3, 2022.

[28] P. Majaranta and A. Bulling, "Eye tracking and eye-based human–computer interaction," *Advances in physiological computing*, pp. 39–65, 2014.

[29] K. Krafka, A. Khosla, P. Kellnhofer, H. Kannan, S. Bhandarkar, W. Matusik, and A. Torralba, "Eye tracking for everyone," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2176–2184.

[30] O. Palinko, A. L. Kun, A. Shyrokov, and P. Heeman, "Estimating cognitive load using remote eye tracking in a driving simulator," in *Proceedings of the 2010 symposium on eye-tracking research & applications*, 2010, pp. 141–144.

[31] M. I. Ahmad, I. Keller, D. A. Robb, and K. S. Lohan, "A framework to estimate cognitive load using physiological data," *Personal and Ubiquitous Computing*, pp. 1–15, 2020.

[32] Y. Guo, D. Freer, F. Deligianni, and G.-Z. Yang, "Eye-tracking for performance evaluation and workload estimation in space telerobotic training," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 1, pp. 1–11, 2022.

[33] R. Gavas, D. Chatterjee, and A. Sinha, "Estimation of cognitive load based on the pupil size dilation," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1499–1504.

[34] P. van der Wel and H. van Steenbergen, "Pupil dilation as an index of effort in cognitive control tasks: A review," *Psychonomic Bulletin & Review*, vol. 25, no. 6, pp. 2005–2015, 2018.

[35] Y. Wang, G. Naylor, S. E. Kramer, A. A. Zekveld, D. Wendt, B. Ohlenforst, and T. Lunner, "Relations between self-reported daily-life fatigue, hearing status, and pupil dilation during a speech perception in noise task," *Ear Hear*, vol. 39, no. 3, pp. 573–582, May/Jun 2018.

[36] P. Azizinezhad, H. Ghonchi, and A. Chowdhury, "Pupil diameter classification using machine learning during human-computer interaction," in *2024 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*. IEEE, 2024, pp. 1–6.