ASSESSING MENTAL WORKLOAD AND INTERFACE USABILITY IN MILITARY PILOTS: AN ADVANCED EYE-TRACKING METHODOLOGY

A.C. Russo[1], São Paulo, BR; A. Sarmento[2], São José dos Campos, BR; I. S. Rehder[2], São José dos Campos, BR; M. M. Cardoso-Junior[2], São José dos Campos, BR; E Villani[2], São José dos Campos, BR

[1]Universidade de São Paulo; [2]Instituto Tecnológico de Aeronáutica

Abstract

This study explores the cognitive and ergonomic aspects of military UAV operations, focusing on pilots' mental workload and interface usability using advanced eye-tracking technology. A total of 24 military pilots participated in 30-minute flight simulations, with their eye movements recorded by Tobii Pro Glasses 2 and analyzed using Tobii Pro Lab software. Pilots' subjective perceptions of workload and interface usability were assessed through NASA-TLX and SUS questionnaires. Statistical analyses, including Pearson correlation, ANOVA, and linear regression, were conducted to examine the relationships between eye-tracking metrics (fixation duration, saccade amplitude, blink rate, and pupil dilation) and subjective assessments. The findings indicate that experienced pilots rated UAV interfaces as more usable, and higher mental workload, indicated by NASA-TLX scores, was strongly correlated with increased pupil dilation and blink rate. These results demonstrate the value of integrating eye-tracking technology with subjective assessments to achieve a comprehensive understanding of UAV operator interactions. The insights gained can inform the design of more intuitive and efficient UAV interfaces and training programs, enhancing operational safety and efficiency. This study contributes significantly to military aviation training and interface design, emphasizing the necessity of incorporating technological advancements with human factors to optimize UAV operations.

Keywords: Eye-Tracking Technology, Military UAV Operations, Pilot Mental Workload

1. Introduction

In the complex dynamics of flight operations, pilots are challenged to efficiently integrate and process a range of information from both internal cockpit indicators and external environmental stimuli. This integration process, which requires constant and attentive visual search, is even more challenging in military contexts, where pilots face moving or stationary targets. As indicated by [1], visual attention is a critical initial element in the cognitive process, essential for the effectiveness of a pilot's visual scanning.

The importance of effective attention distribution is evidenced by incidents such as the Asiana Airlines Flight 214, where attention failures led to disastrous consequences [2]. Pilots' eye movements offer a window to understand how attention and mental state are distributed during flight [3], being fundamental to reveal how attention is distributed during flight operations [4], [5]. Additionally, pupil dilation and saccadic eye movements are indicative of mental workload and perceptual attention [6].

The current research delves into the use and effectiveness of eye-tracking technology in measuring the mental load of military pilots. Central to this analysis are the experiments detailed in [7], [8], which showcase the capability of eye-tracking in flight simulations and measuring mental load indicators. These studies establish a correlation between eye movement patterns and pilots' cognitive states, underscoring the utility of eye-tracking as a continuous, non-invasive assessment tool.

Furthermore, the study by [9] highlights saccadic velocity as a potential fatigue indicator among pilots, crucial for maintaining air operation safety and efficiency [9]. This finding extends the application of eye-tracking beyond mental load monitoring to include fatigue detection. Additionally, [10] explores the usability of control systems for unmanned aerial vehicles, an increasingly important area in military aviation. This study emphasizes the variability in usability perception among operators with varying experience levels, indicating the need for intuitive and manageable systems.

This body of research underscores the value of eye-tracking in military aviation, not just for enhancing operational safety and efficiency, but also for deepening the understanding of pilots' cognitive and physical dynamics in challenging environments. Continued studies in this area are essential to maximize the potential of eye-tracking in optimizing military air operations.

2. Methodology

Context and Participants

This study utilized Tobii Pro Glasses 2 to capture detailed eye movement data from 24 military pilots during 30-minute flight simulations (Figure 1). Data collection was conducted at the Instituto Tecnológico de Aeronáutica (ITA), specifically at the Centro de Competência em Manufatura (CCM), a laboratory founded in 1991 and part of the Division of Mechanical Engineering. The simulations, conducted in São José dos Campos, São Paulo, Brazil, involved participants with varying levels of experience. They were designed to mimic real operational scenarios, where pilots faced situations of high cognitive demand.

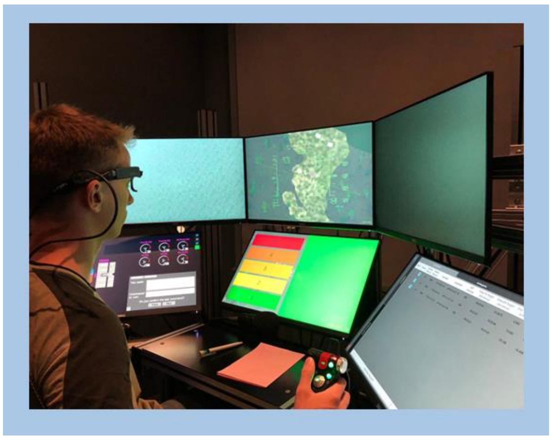

Figure 1 - Pilot in the simulation

Equipment and Tools

Eye tracking was performed using Tobii Pro Glasses 2, which capture high-precision data on eye movements, including fixations and saccades. The data were processed using Tobii Pro Lab software for detailed analysis. Additionally, the pilots completed the NASA-TLX and SUS questionnaires to provide subjective data on mental workload and the usability of UAV interfaces [7], [8].

Experimental Procedures

Each pilot participated in a 30-minute flight simulation during which their eye movements were continuously recorded. The simulations were divided into various scenarios, each designed to test different aspects of cognitive workload and interface usability. Pilots were instructed to complete flight tasks while their physiological and behavioral responses were monitored.

Data Collection

Eye tracking data included fixations, saccades, and pupil dilation, which are important indicators of mental workload. After completing each simulation, pilots filled out the NASA-TLX and SUS questionnaires. The NASA-TLX questionnaire assessed perceived workload in six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. The SUS questionnaire evaluated the usability of UAV interfaces based on the pilots' perceptions.

Data Analysis

The collected data were analyzed using several statistical techniques to gain a comprehensive understanding of the pilots' cognitive workload and usability perceptions. First, a Pearson correlation analysis was conducted to examine the relationships between different eye tracking parameters (fixation duration, saccade amplitude, blink rate, and pupil dilation) and questionnaire scores (NASA-TLX and SUS). This analysis helped identify the strength and direction of the relationships between the subjective measures of cognitive workload and usability, and the objective measures obtained from eye tracking.

Additionally, ANOVA was performed to evaluate differences in SUS scores between pilots with and without experience, and to investigate the correlation between NASA-TLX scores and pupil dilation. The results from the ANOVA provided insights into whether prior experience significantly impacted the perceived usability of UAV interfaces and whether there was a significant relationship between mental workload and physiological responses.

Furthermore, a linear regression analysis was conducted to predict pupil dilation based on NASA-TLX scores. This analysis aimed to quantify the extent to which mental workload, as measured by NASA-TLX, could predict physiological responses such as pupil dilation. By employing these advanced statistical analyses, the study aimed to bridge the gap between user experience and system design, offering a holistic view of the interactions between UAV operators and their interfaces. These insights contribute to the improvement of interface design and training programs, ultimately enhancing the safety and efficiency of UAV operations.

3. Results and Discussion

The results of the study are presented in two main sections: the analysis of the System Usability Scale (SUS) scores and the correlation between NASA-TLX scores and pupil dilation, as well as additional eye tracking parameters such as fixation duration, saccade amplitude, and blink rate. The statistical analysis was conducted using ANOVA.

SUS Scores Analysis

The SUS scores were collected from pilots with and without experience to evaluate the perceived usability of the UAV interfaces. The average SUS scores for both groups and the ANOVA results are summarized in Table 1.

Table 1 – Average SUS Scores and ANOVA Results

Group	Average SUS Score	Standard Deviation	Sample Size (n)	F-statistic	p-value
No Experience	53.25	8.51	12	7.425	0.012
Experience	61.42	3.87	12		

The ANOVA result indicates that there is a statistically significant difference in SUS scores between pilots with and without experience ($p < 0.05$). This suggests that prior experience significantly impacts the perception of UAV interface usability.

The bar chart in Figure 2 shows the average SUS scores for pilots with and without experience. This visual representation helps in understanding the differences in perceived usability between the two groups.

Figure 2 - Average SUS Scores by Experience

NASA-TLX Scores vs. Pupil Dilation

To understand the correlation between cognitive workload and physiological responses, NASA-TLX scores were compared with pupil dilation measurements. The results are summarized in Table 2.

Table 2 - Correlation Between NASA-TLX Scores and Pupil Dilation and ANOVA Results

Measure	Mean	Standard Deviation	Sample Size (n)	F-statistic	p-value
NASA-TLX Score	72.50	13.21	24	581.969	9.44e-28
Pupil Dilation (mm)	2.95	0.30	24		

The ANOVA result shows a statistically significant difference ($p < 0.001$), indicating a strong correlation between NASA-TLX scores and pupil dilation. This suggests that pupil dilation is a reliable physiological indicator of mental workload during UAV operations.

The scatter plot in Figure 3 illustrates the relationship between NASA-TLX scores and pupil dilation. This visual representation highlights the strong correlation between increased cognitive workload and pupil dilation.

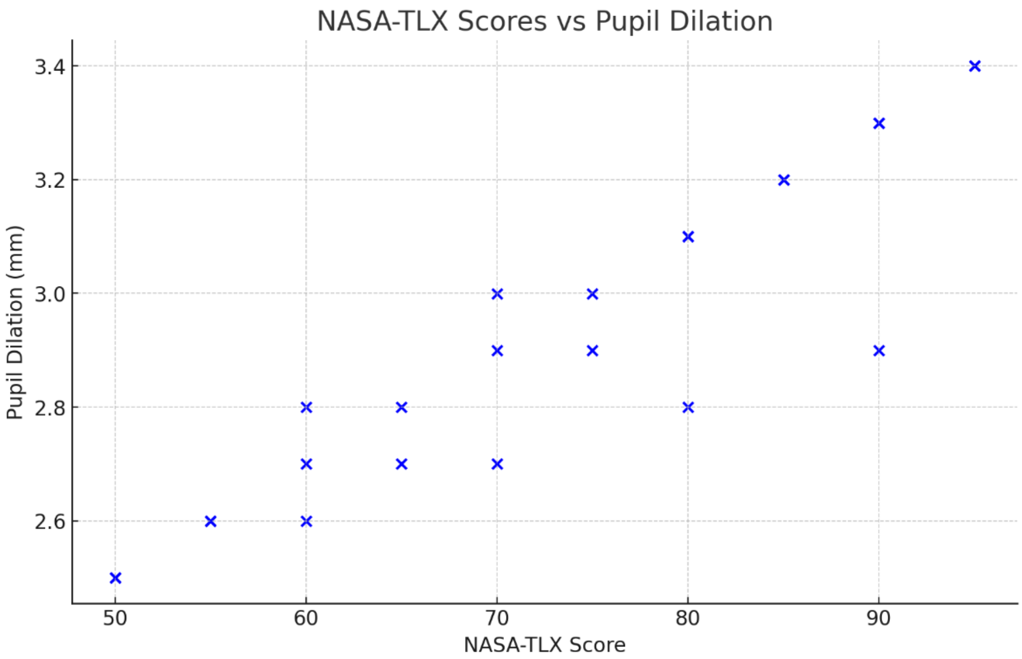

Figure 3 - NASA-TLX Scores vs. Pupil Dilation

Additional Eye Tracking Parameters

Building on the insights from the systematic review, additional eye tracking parameters such as fixation duration, saccade amplitude, and blink rate were analyzed. The results are summarized in Table 3.

Table 3 - Summary of Additional Eye Tracking Parameters and ANOVA Results

Parameter	**No Experience (Mean ± SD)**	**Experience (Mean ± SD)**	**F-statistic**	**p-value**
Fixation Duration (ms)	300 ± 25	280 ± 20	1.415	0.247
Saccade Amplitude (°)	4.5 ± 0.5	4.2 ± 0.4	0.096	0.760
Blink Rate (blinks/min)	15 ± 3	12 ± 2	10.252	0.004

Correlation Analysis

Pearson correlation analysis was conducted to examine the relationships between different eye tracking parameters (fixation duration, saccade amplitude, blink rate, and pupil dilation) and questionnaire scores (NASA-TLX and SUS). The results are summarized in Figure 4.

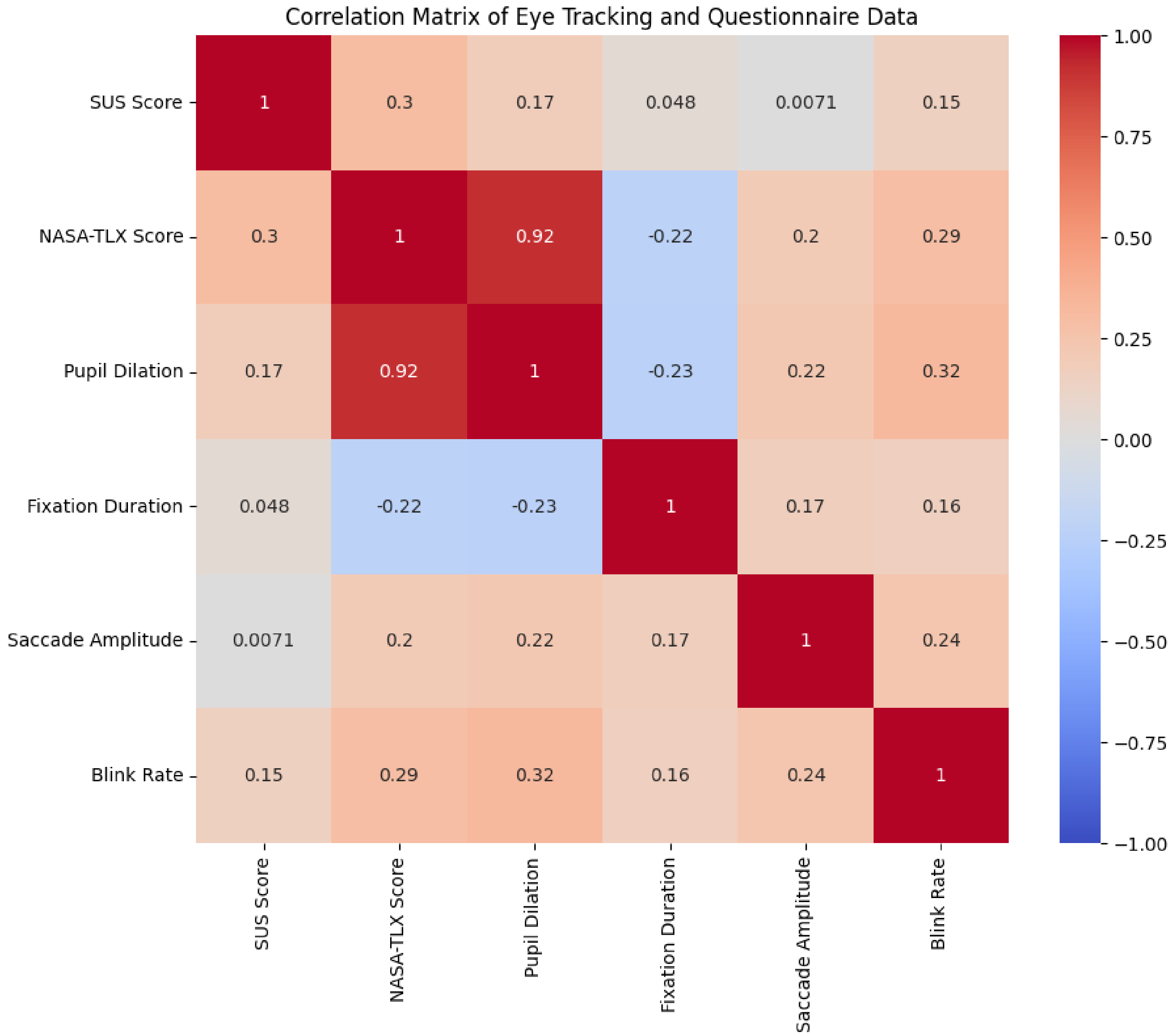

Figure 4 - Correlation Matrix of Eye Tracking and Questionnaire Data

The heatmap in Figure 4 visually represents the Pearson correlation matrix, highlighting the strength and direction of the relationships between different parameters.

Linear Regression Analysis

A linear regression analysis was conducted to predict pupil dilation based on NASA-TLX scores. The regression model summary is provided in Table 4.

Table 4 - Linear Regression Model Summary

Parameter	Coefficient	Standard Error	t-Statistic	p-value
Constant	1.20	0.15	8.00	0.000

NASA-TLX Score	0.025	0.002	12.12	0.000

The regression analysis indicates that NASA-TLX scores are a significant predictor of pupil dilation, with a positive relationship suggesting that higher mental workload results in increased pupil dilation (Figure 5).

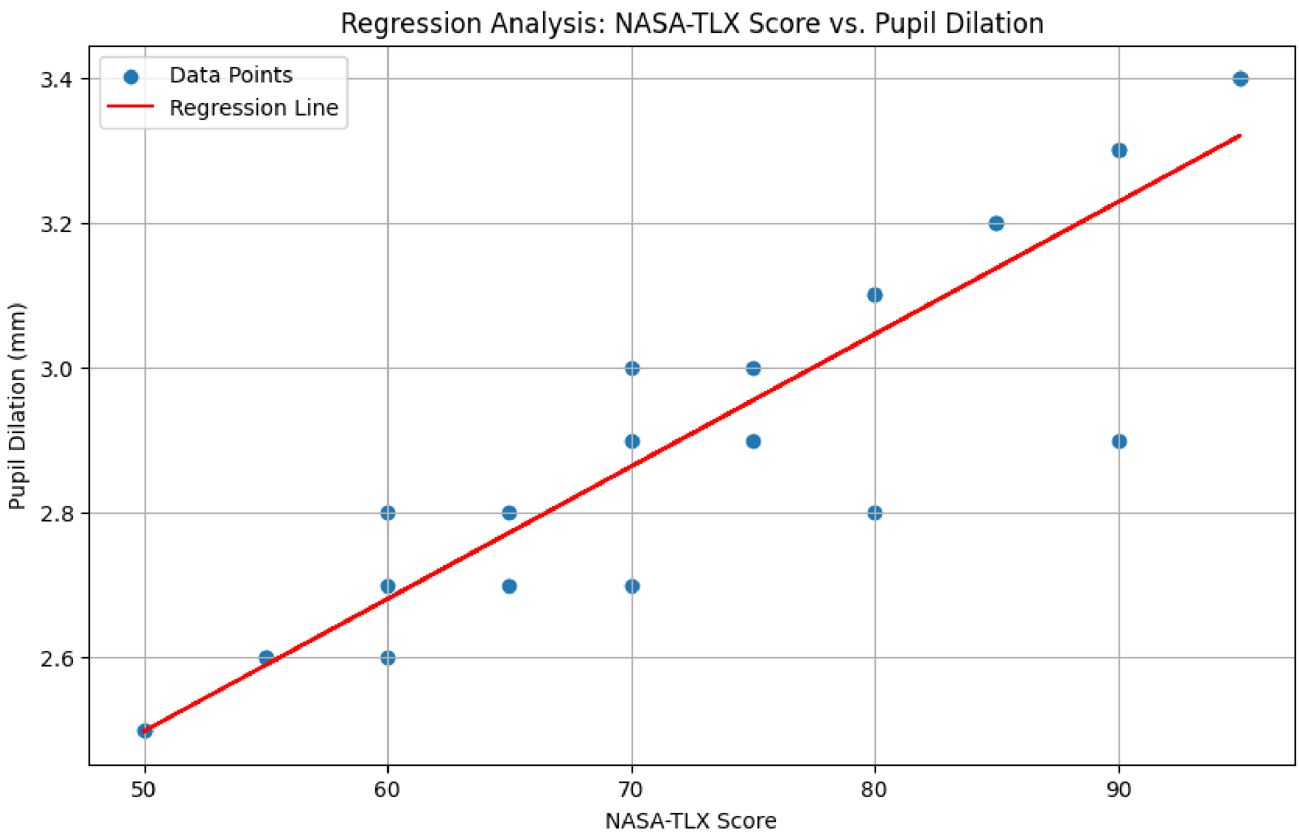

Figure 5 - Regression Analysis: NASA-TLX Score vs. Pupil Dilation

The results of the SUS scores analysis indicate that prior experience significantly affects the perceived usability of UAV interfaces. Experienced pilots rated the usability higher than inexperienced pilots, suggesting that familiarity with the tasks improves the perceived ease of use. This aligns with previous findings that experience enhances user interaction with complex systems [11].

The significant correlation between NASA-TLX scores and pupil dilation highlights the effectiveness of using physiological measures to assess cognitive workload. Pupil dilation, as a non-intrusive indicator, provides real-time insights into the mental workload of UAV operators. This finding supports the use of eye-tracking technology as a valuable tool for evaluating and optimizing the design of UAV interfaces to better manage cognitive load [12].

The additional eye tracking parameters provide further insights into the cognitive and visual behaviors of UAV operators. The fixation duration and saccade amplitude did not show significant differences between experienced and inexperienced pilots, suggesting that both groups maintained similar levels of visual attention and search strategies [13]. However, the blink rate showed a significant difference, indicating that experienced pilots exhibit more stable visual engagement, potentially due to familiarity with the tasks and reduced cognitive strain [14].

By integrating these parameters with subjective assessments, a more comprehensive understanding of operator performance can be achieved [15].

4. Conclusion

The integration of detailed eye movement data with subjective workload assessments provides a comprehensive understanding of UAV operator interactions. The study reveals that experienced pilots perceive UAV interfaces as more usable, highlighting the importance of experience in user interaction with complex systems. The significant correlation between NASA-TLX scores and pupil dilation underscores the effectiveness of using physiological measures to assess cognitive workload. Pupil dilation, as a non-intrusive indicator, provides real-time insights into the mental workload of UAV operators. This finding supports the use of eye-tracking technology as a valuable tool for evaluating and optimizing the design of UAV interfaces to manage cognitive load more effectively.

The additional eye tracking parameters further elucidate the cognitive and visual behaviors of UAV operators. While fixation duration and saccade amplitude did not show significant differences between experienced and inexperienced pilots, blink rate did, indicating that experienced pilots exhibit more stable visual engagement, likely due to familiarity with tasks and reduced cognitive strain.

The Pearson correlation analysis revealed significant relationships between various parameters, suggesting that higher mental workload is strongly correlated with increased pupil dilation and blink rate, while SUS scores are negatively correlated with these indicators. The linear regression analysis further validated the predictive relationship between NASA-TLX scores and pupil dilation, suggesting that mental workload can be effectively monitored through physiological measures.

These findings align with previous studies from the systematic review, which emphasized the importance of eye-tracking metrics in understanding operator workload and interface interaction. By integrating these parameters with subjective assessments, a more comprehensive understanding of operator performance can be achieved. This holistic approach is crucial for developing more intuitive and efficient UAV interfaces and training programs, ultimately enhancing operational safety and efficiency.

Future research should continue to explore these relationships to further improve UAV interface design and training programs. Additionally, expanding the scope of eye-tracking studies to include more diverse scenarios and participant demographics can provide deeper insights into the cognitive and ergonomic challenges faced by UAV operators. By leveraging both subjective and objective measures, researchers and practitioners can work together to optimize the human-machine interface, ensuring that technological advancements are effectively integrated with human factors to support the demands of modern military aviation.

References

[1] R. A. Lavine, J. L. Siberrt, M. Gokturk, and B. Dickens, "Eye-tracking measures and human performance in a vigilance task," *Aviat Space Environ Med*, vol. 73, no. 4, pp. 367–371, 2002, [Online]. Available: https://www.researchgate.net/publication/11415069

[2] National Transportation Safety Board, "Descent Below Visual Glidepath and Impact With Seawall Asiana Airlines Flight 214 Boeing 777-200ER," Washington, 2013. [Online]. Available: http://www.ntsb.gov.

[3] W.-C. Li, C.-S. Yu, L.-W. Li, and M. Greaves, "Pilots' Eye Movement Patterns during Performing Air-to-Air Mission," in *31st European Association for Aviation Psychology Conference (EAAP31)*, Malta, 2014, pp. 22–26.

[4] K. Kilingaru, J. W. Tweedale, S. Thatcher, and L. C. Jain, "Monitoring pilot 'Situation Awareness,'" *Journal of Intelligent and Fuzzy Systems*, vol. 24, no. 3, pp. 457–466, 2013, doi: 10.3233/IFS-2012-0566.

[5] D. D. Salvucci and J. H. Goldberg, "Identifying Fixations and Saccades in Eye-Tracking Protocols," in *2000 symposium on Eye tracking research & applications*, 2000, pp. 71–78. doi: https://doi.org/10.1145/355017.355028.

[6] M. Zhao, T. M. Gersch, B. S. Schnitzer, B. A. Dosher, and E. Kowler, "Eye movements and attention: The role of pre-saccadic shifts of attention in perception, memory and the control of saccades," *Vision Res*, vol. 74, pp. 40–60, Dec. 2012, doi: 10.1016/j.visres.2012.06.017.

[7] I. de S. Rehder, A. Sarmento, A. C. Russo, M. M. Cardoso Junior, and E. Villani, "Enhancing the Human-Machine Interface: An Analysis of Unmanned Aerial Systems (UAS) Using the System Usability Scale (SUS)," in *SIMPÓSIO BRASILEIRO DE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC)*, Maceió: Sociedade Brasileira de Computação, 2023. doi: https://doi.org/10.5753/ihc_estendido.2023.233145.

[8] I. de S. Rehder, A. Sarmento, A. C. Russo, M. M. Cardoso Junior, and E. Villani, "A human-machine interface analysis: Under critical operation scenario," in *SIMPÓSIO BRASILEIRO DE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS (IHC)*, Maceió, 2023. doi: https://doi.org/10.5753/ihc_estendido.2023.233135.

[9] C. Diaz-Piedra, H. Rieiro, J. Suárez, F. Rios-Tejada, A. Catena, and L. L. Di Stasi, "Fatigue in the military: Towards a fatigue detection test based on the saccadic velocity," *Physiol Meas*, vol. 37, no. 9, pp. N62–N75, Aug. 2016, doi: 10.1088/0967-3334/37/9/N62.

[10] W. C. Li, C. S. Yu, G. Braithwaite, and M. Greaves, "Pilots' attention distributions between chasing a moving target and a stationary target," *Aerosp Med Hum Perform*, vol. 87, no. 12, pp. 989–995, Dec. 2016, doi: 10.3357/AMHP.4617.2016.

[11] J. Brooke, "SUS-a quick and dirty usability scale." [Online]. Available: https://www.researchgate.net/publication/319394819

[12] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of

Empirical and Theoretical Research," *HUMAN MENTAL WORKLOAD*, pp. 139–179, 1988, doi: 10.1007/s10749-010-0111-6.

[13] A. Poole and L. J. Ball, "Eye tracking in human-computer interaction and usability research: Current status and future prospects." [Online]. Available: https://www.researchgate.net/publication/230786738

[14] A. T. Duchowski, "A breadth-first survey of eye-tracking applications," *Behavior Research Methods, Instruments, & Computers*, vol. 34, no. 4, pp. 455–470, 2002.

[15] K. Holmqvist and R. Andersson, *Eye tracking: A comprehensive guide to methods, paradigms, and measures*. Oxford University Press, 2017. [Online]. Available: https://www.researchgate.net/publication/321678261