Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira, Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria

doi: 10.3850/ESREL2026061419_esrel26-p26325-cd

The Eyes Tell The Load: Saccadic Intrusions as Indicators of Cognitive Demand in Latency

Marina Ronconi de Oliveira
Aeronautics Institute of Technology, Brazil E-mail: marina.oliveira@ccm-ita.org.br

Marina Laís Rosa
Aeronautics Institute of Technology, Brazil E-mail: marina.rosa@ccm-ita.org.br

Ivan de Souza Rehder
Aeronautics Institute of Technology, Brazil E-mail: iva.rehder@ccm-ita.org.br

Emilia Villani
Aeronautics Institute of Technology, Brazil E-mail: emilia.villani@ccm-ita.org.br

Moacyr Machado Cardoso Junior
Aeronautics Institute of Technology, Brazil E-mail: moacyr.cardoso@gp.ita.br

This study investigates the correlation between mental workload and saccadic intrusions recorded through eye tracking during the landing of a remotely piloted aircraft (or also called as Unmanned Aerial Vehicles UAVs) under communication delay conditions. The experiment simulates a satellite link operation with a 2-second delay between the pilot's command and the aircraft's response, a condition that significantly increases task complexity and cognitive demands. Three configurations of an adaptive Head-Up Display (HUD) were evaluated to identify which one best supports pilot performance under latency conditions. Objective measures of saccadic intrusions were correlated with subjective assessments of mental workload obtained through the NASA-TLX, SWORD, and Instantaneous Self-Assessment (ISA) questionnaires. The results indicated a convergence among the subjective metrics ($r > 0.98$). A strong negative correlation was observed between perceived workload and the frequency of saccadic intrusions ($r = -0.88$), providing evidence of the cortical suppression phenomenon under high visual demand. The 30-second temporal analysis confirmed that the No-HUD condition imposes a statistically significant state of attentional tunneling ($\rho = -0.56; p = 0.0456$), whereas the HUD interfaces mitigated this oculomotor rigidity, allowing more natural and less fatiguing visual processing.

Keywords: Mental workload, Eye tracking, Saccadic intrusions, Human-machine interface, Subjective questionnaires.

1. Introduction

The use of Unmanned Aerial Vehicles (UAVs) is growing rapidly in both civilian and military fields, especially in high-risk operations. These systems allow humans to work in hard-to-reach or dangerous places without putting pilots in physical danger. However, this change deeply affects how flights are conducted (McCarley and Wickens, 2005). Unlike traditional aviation, where the pilot is inside the aircraft and uses all their senses to feel the environment (Landry, 2018), remote pilots interact with the vehicle through digital screens and communication links. In this situation, the pilot's awareness depends almost entirely on visual information on a screen, which creates new challenges for Human Factors (Gawron, 1998).

Some long-endurance UAVs operate far from the control station, requiring communications ”beyond line of sight” (BLOS) (McCarley and Wickens, 2005). These signals often have delays or latencies of one second or more. Any delay longer than one second makes real-time manual control very difficult (Wickens and Hollands, 2000). In these cases, the immediate link between an action and its effect is broken. As a result, the operator must mentally predict what the aircraft will do and deal with more uncertainty, which greatly increases their mental workload. During critical phases like approach and landing, where small corrections are vital, this extra effort can hurt both performance and safety.

To help remote pilots, different technologies have been suggested, such as higher levels of automation and advanced interfaces like Head-Up Displays (HUDs) (Caldwell, 2003). In theory, these tools help the pilot understand the situation by highlighting important information. However, if they are poorly designed or too crowded, they can distract the operator and increase mental demand, causing more harm than good.

Because of this, it is essential to have reliable ways to measure mental workload. Subjective scales like NASA-TLX (Hart and Staveland, 1988), SWORD (Vidulich and Tsang, 1987), and the ISA (Instantaneous Self-Assessment), Tattersall and Foord, 1996) are commonly used to capture how much effort an operator feels they are using. While useful, these methods can be biased by personal opinion, which is why researchers are looking for more objective indicators.

Eye-tracking has become a promising tool for studying how the brain processes information during complex tasks and how eye movements indicate where attention is focused. Specifically, saccadic intrusions, small, involuntary eye movements that interrupt focus, have been linked to increased mental effort and cognitive activation (Tokuda et al., 2009).

This study explores this environment. By analyzing saccadic intrusions during cruise flight and landing, with a simulated 2-second delay, and comparing them with subjective workload scores (NASA-TLX, SWORD, and ISA), this research aims to understand how mental demand affects eye patterns and how different HUD designs influence the pilot in complex situations.

2. Literature Review

The eye conveys about individuals' mental state during task performance. Features such as pupil size and eye movement patterns have been shown to reflect mental workload in previous studies conducted by Tokuda et al. (2009), Babu et al. (2019) and Biswas and Prabhakar (2018). Among those eye-related features, eye movements are notable due to their robustness to environmental factors. In daily life, ocular trajectories switch between different types of movement to navigate efficiently through the environment, namely saccades and fixations. Saccades are gaze-shifting movements, scanning the environment in a search for information. They are brief events, typically lasting from 20 to 200ms, during which the point of fixation can shift at a velocity up to 50 degrees per second. Through fixation periods, which exceed 200 ms, the eyes are stable at a focus point, analysing the target aiming to understand it and capture information. Although stable, the eyes are not motionless. According to Abadi and Gowen (2004), even during fixation, there are slight deviations from the focus point. Slow drifts, microsaccades, and saccadic intrusions are present in this type of movement, each with its own characteristics. Slow drifts are small angular movements, whose duration can reach just a few seconds. Microsaccades, as regular saccades, are related to gaze re-orientation, but with an almost unnoticeable change of gaze point (Hafed and Clark, 2002), as they have a maximum amplitude of $0.2°$.

Between those movements, the noteworthy one for this article is the saccadic intrusions (SI). Those are involuntary variations from fixation movements, they are characterized for deviating from the gaze point and quickly returning to it. These oscillations occur mainly in the horizontal axis and are coordinated by both eyes (Abadi and Gowen, 2004). When compared to other slighter deviations, SI has the largest amplitude (Tokuda et al., 2011), which range is $0.1°$ - $4.1°$, its duration starts in 20 ms and can reach 870 ms. As an eye movement, the saccadic intrusions were also correlated with cognitive workload indexes. Features such as the number of occurrences and velocity have been shown to have a positive relation; as the frequency of occurrence or velocity increases, the workload is higher (Tokuda et al., 2011).

Complementing this physiological perspective, Shaffer et al. (2003) provides a neurological basis for these variations through the concept of cortical suppression. Their research indicates that under high visual task demands, cortical attentional systems may suppress inappropriate saccades to prevent the eyes from diverting from objects of interest. Therefore, the modulation of these intrusive movements serves as an objective indicator of the competition for attentional resources, where more demanding conditions—such as those involving unpredictable targets or complex maneuvers—force the oculomotor system to prioritize stability to maintain performance.

Despite the foundational nature of these studies, the literature on saccadic intrusions as a workload metric has seen a period of limited empirical expansion, with core models established over a decade ago. However, the emergence of modern complex systems, such as UAVs operations, creates a renewed demand for these metrics. There is a notable gap in understanding how SI patterns behave under contemporary challenges like satellite communication latency and adaptive digital interfaces. Therefore, this study aims to bridge this gap by applying these established physiological markers to a high-stakes aeronautical context, updating the relevance of SI as a robust tool for real-time cognitive monitoring in current flight simulation technologies.

3. Materials and Methods

This study investigates the correlation between mental workload and saccadic intrusions recorded through eye-tracking during the landing of a UAV. Objective measures of saccadic intrusions were correlated with subjective assessments from NASA-TLX, SWORD, and ISA questionnaires. The research was conducted at the Aeronautical Institute of Technology (Instituto Tecnológico de Aeronáutica) (ITA), within the Competence Center of Manufacturing (Centro de Competência em Manufatura) (CCM).

4. The Scenario

This experiment is part of a larger initiative called the Air Domain Study (ADS), which follows a ”design–test–analysis” protocol. The flights were conducted using the ADS Simulator (Sarmento et al., 2022), a computational platform designed to replicate aircraft behavior in a controlled environment. The tests ran in real-time using Simulink® and FlightGear®, while the cockpit interface was rendered using Unity®.

The scenario involved a simulated UAV flight between two cities. During the flight, a critical failure was introduced, forcing the pilot to perform an emergency landing at an unmapped airfield. Because of this failure, the automatic landing configuration system became unavailable, requiring the operator to take full manual control of the aircraft.

To increase the cognitive demand, a fixed two-second satellite communication latency was applied. This created a delay between the pilot’s commands and the aircraft’s response.

Each session consisted of a visual approach to the São José dos Campos airport, starting from cruise altitude. The only factor changed during the tests was the HUD configuration, allowing us to analyze how different interfaces help the pilot manage the communication delay.

4.1. *HUDs*

In the No HUD setup [Figure 1], the pilots could only see the Primary Flight Display (PFD). All flight data — such as heading, speed, altitude, and pitch — had a 2-second delay, meaning the pilots did not see the aircraft’s movement in real-time.

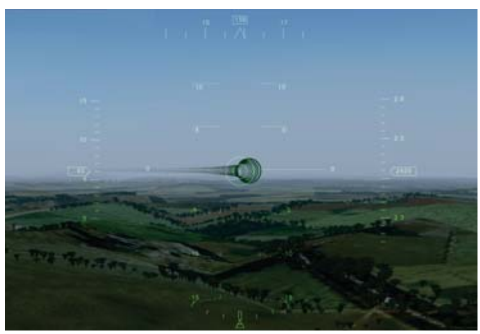

Fig. 1.: No HUD.

The setup called HUD 1 [Figure 2] added digital information over the forward view. It included numeric readouts for bank angle, roll rate, pitch angle, and pitch rate. To help with the delay, it also showed a predicted flight path represented by a series of blue squares. These squares showed where the aircraft would be, calculated by the onboard computer.

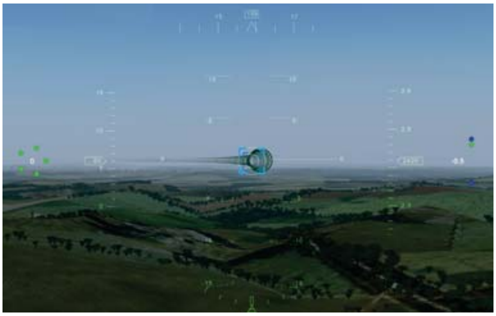

Fig. 2.: HUD 1.

And in the setup HUD 2 [Figure 3], this version improved HUD 1 by adding graphical tapes for easier reading of the bank, roll, and pitch data. Instead of just squares, it used a magenta ”flight-path ball” to show the predicted path. It also included a throttle status indicator and offered two modes (Cruise and Landing) that adjusted the information based on the flight phase.

Fig. 3.: HUD 2.

4.2. *The Experiment*

The experiment was conducted with 24 volunteers, all of whom were military pilots from the Brazilian Air Force (FAB). While all participants were experienced aviators, prior experience with UAVs was not a requirement for inclusion. All participants were informed about the study's objectives and signed the Informed Consent Form, following ethical research standards (CAAE 77429824.7.0000.5503). Before the data collection sessions, participants underwent a standardized training phase using a one-second communication delay to become familiar with the simulation environment and the specific dynamics of latent commands.

During data analysis, it was found that 4 eye-tracking recordings were corrupted due to intermittent hardware connectivity issues that prevented the complete storage of gaze data. Therefore, the final analysis was conducted with 18 participants. This number was chosen to maintain a balanced experimental design. Since there are 3 HUD configurations, the presentation order was randomized across 6 possible sequences to control for order and learning effects. Using 18 participants allowed for exactly 3 complete repetitions of these 6 sequences, ensuring the results are statistically consistent.

5. Data Analysis

5.1. *Data Preprocessing*

Aiming to identify saccadic intrusions, an algorithm was developed based on previous studies by Liu et al. (2022) and Tokuda et al. (2009). The analysis was conducted using eye-tracking data collected from the Tobii Pro Glasses 2 were utilized, which operates at a sampling rate of 50 Hz. The algorithm development and implementation relied on data pre-processing followed by its input into the system.

Saccadic intrusion is defined as a conjugate eye movement, which leads to a similar number of occurrences in both eyes (Tokuda et al., 2011), therefore its identification was executed using data from a single eye. Following the data cleaning and organization, essential features of gaze behavior were highlighted, namely gaze point, pupil position and fixation point. All extracted eye-tracking information was grounded in the same coordinate system: Head Unit Coordinate System (HUCS), whose origin is at the scene camera center, positioned between the user's eyebrows. Although the values were initially established in this coordinate system, detailed data about individual eye trajectory during the flights could not be obtained. As a consequence, a transformation of the origin point was required, relocating it to the right eye center when the gaze direction is towards the screen's central point.

The initial procedure to shift the coordinate system reference was to filter the points of fixation, underlining those towards the display center. Subsequently, the mean of right pupil position associated with those fixation points obtained, resulting in a coordinate pair that defined the new reference point of the right eye coordinate system (RE). The gaze point translation to the new coordinate system was based on the equation Equation 1:

$$\text{Gaze Point}_{RE} = \text{Gaze Point}_{HUCS} - \text{Origin}_{RE} \tag{1}$$

SI's definition proposed by Tokuda et al. (2009) emphasizes the angular trajectory of the eye for its identification. Thus, after correcting the coordinate system, a calculation was performed to calculate the eye angle (θ) through the X and Z axes, considering the horizontal aspect of these movements, see equation Equation 2:

$$\theta = \text{atan2}\left(\frac{X_{GazePoint\_RE}}{Z_{GazePoint\_RE}}\right) \times \left(\frac{180}{\pi}\right) \tag{2}$$

The mathematical operation used to estimate the vector angle of the gaze point accounts for points located in the second quadrant of the RE coordinate system. Using the two-argument arctangent function, it was possible to determine the quadrant and the absolute eye angle. For standardization and ease in building graphics, results of this operation were normalized with the maximum value reachable by the eye, which was defined as 180 degrees.

5.2. *The Algorithm*

The Python code developed relies on conditional logic to classify the eye-tracking samples according to the definition of a saccadic intrusion. Analysis is based on the eye angle, wherein the algorithm computes, in a 2000ms moving window, the median of the normalized angle and defines it as the baseline, corresponding to a trajectory without abrupt oscillations. Thereafter, the instant sample angular position is compared to the baseline. When deviation is in the predefined SI thresholds, the algorithm counts the number of consecutives samples in the same pattern. Once this pattern is disrupted, the algorithm verifies if the duration consists of a saccadic intrusion. As a result, this algorithm returns multiple movement features, including frequency, velocity, duration and amplitude. In addition to these measures, it also plots the eye trajectory during the selected period, as shown in figure 4.

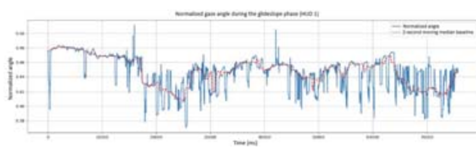

Fig. 4.: Normalized gaze angle during the glideslope phase using the HUD 1 configuration. The red dashed line represents the 2-second moving median baseline used to identify saccadic intrusions.

6. Results

6.1. *Convergence of Subjective Metrics*

The initial data analysis revealed an exceptional level of consistency among the self-assessment instruments employed. An extremely strong Pearson correlation (see Table 1) was observed between the Instantaneous Self-Assessment (ISA) and the NASA Task Load Index (NASA-TLX) ($r = 0.985$; $p = 0.109$), as well as between the ISA and the Subjective Workload Dominance (SWORD) technique ($r = 0.983$; $p = 0.118$). The correlation between the post-flight scales NASA-TLX and SWORD was perfect ($r = 1.00$; $p < 0.001$), demonstrating an absolute convergence between these metrics (Figure 5). Although the p-values associated with the correlations involving the ISA were slightly above the conventional 0.05 threshold due to the statistical stringency imposed by the small sample size of the means ($n = 3$), the magnitude of the coefficients ($r > 0.98$) provides strong validation for the mental workload construct investigated, indicating that the measures captured the same underlying variation in cognitive effort experienced by the pilots.

Variable	p-value	Result
intrusion_amplitude	0.0928	Normal (Pearson)
intrusion_count	0.106	Normal (Pearson)
intrusion_duration	0.129	Normal (Pearson)
intrusion_speed	0.591	Normal (Pearson)
isa_score	0.921	Normal (Pearson)
nasa	0.716	Normal (Pearson)
sword	0.659	Normal (Pearson)

Table 1.: Results of the Pearson normality test for all analyzed variables.

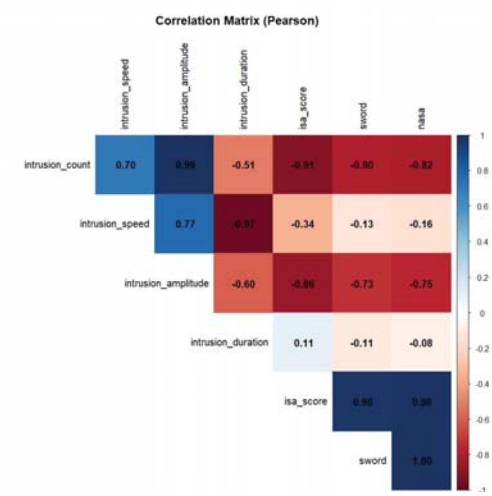

Fig. 5.: Pearson correlation matrix between subjective workload metrics (ISA, SWORD, and NASA-TLX) and saccadic intrusion parameters.

Unlike studies employing pure auditory memory tasks—where saccadic intrusions tend to increase with workload ($r \approx 0.45$, Tokuda et al., 2011), the results of this flight revealed a strong negative correlation between perceived workload and intrusion count ($r = -0.82$ for NASA-TLX; $r = -0.91$ for ISA). This finding confirms the cortical suppression phenomenon documented by Shaffer et al. (2003). Under high visual demand requiring continuous foveation on critical targets, the attentional system inhibits involuntary eye movements to maintain gaze on essential information. Accordingly, a higher number of intrusions during moments reported as “easier” reflects relaxed ocular control, whereas the marked reduction of SIs under high workload serves as a biomarker of visual stabilization required for operational performance.

6.2. *Temporal Analysis and Impact of HUD Interfaces*

To assess the ergonomic effectiveness of the Head-Up Display (HUD) interfaces, a Spearman correlation analysis was conducted using 30-second temporal windows, synchronizing the ISA prompts with the frequency of saccadic intrusions. The results revealed critical distinctions across flight conditions:

No-HUD condition (significance confirmed): This configuration exhibited the strongest and statistically significant negative correlation ($\rho = -0.56; p = 0.0456$). As the p-value is below the 0.05 threshold, this result confirms that, in the absence of projected visual aids, increasing workload significantly reduced the number of intrusions. This pattern suggests a state of severe attentional tunneling, in which the complexity of the conventional cockpit display forced the pilot to exert rigid inhibitory control over gaze in order to process the available information.

HUD 1 and HUD 2 conditions (non-significant trends): Under the HUD conditions, the correlations became weaker and lost statistical significance (HUD 2: $\rho = -0.28; p = 0.3510$; HUD 1: $\rho = 0.26; p = 0.3877$). The lack of significance with the same 13 data points indicates that the adaptive interfaces altered oculomotor dynamics, effectively “breaking” the direct tunneling relationship observed in the control condition.

The superiority of intrusion count as a predictor of workload ($r = -0.91$) compared with intrusion velocity ($r = -0.34$) (Figure 5) reinforces that the frequency of these biological events is a more sensitive indicator for monitoring real-time fluctuations in workload. The fact that the HUD 2 design mitigated the rigidity of oculomotor control

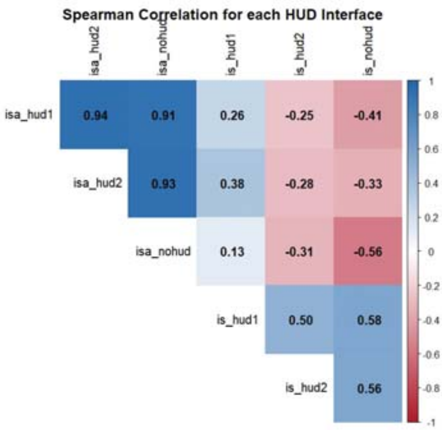

Fig. 6.: Spearman Correlation for each HUD Interface.

(by reducing the magnitude of the negative correlation relative to the No-HUD condition) provides evidence that this interface promotes more natural and less fatiguing visual processing, thereby increasing pilot resilience in high-complexity scenarios.

7. Conclusion

This study demonstrated that the integration of subjective metrics and biometric eye-tracking indicators provides a robust methodology for assessing mental workload in high-fidelity aviation environments. The exceptional convergence among NASA-TLX, SWORD, and ISA ($r > 0.98$) validates the pilots' conscious perception of effort, whereas the analysis of saccadic intrusions adds an objective biological layer of validation, independent of environmental factors such as luminance.

The central finding of this research was the identification of a strong negative correlation between workload and saccadic intrusion count, confirming the occurrence of cortical suppression and attentional tunneling. The reduction in SI frequency under high workload serves as a biomarker of the visual stabilization required for operational performance. Under high visual demand, the attentional system inhibits these involuntary movements to prevent the gaze from diverting from critical targets. Therefore, in this context, fewer intrusions indicate high task engagement and cortical effort, rather than disengagement.

rom a statistical perspective, the No-HUD condition was the only one to reach significance ($p < 0.05$), indicating that the absence of projected visual support forces the pilot to exert rigid and fatiguing oculomotor control in order to process information from the conventional cockpit display. In contrast, the loss of significance under the HUD conditions —particularly with HUD 2 — suggests that the interface design was effective in promoting more varied and natural visual behavior, thereby mitigating the biological effort required for gaze stabilization.

Future developments in adaptive cockpits may exploit the decrease in saccadic intrusion frequency as a biometric trigger for proactive interventions, such as display simplification or automatic task reallocation, thereby ensuring operational safety and preserving the pilot's cognitive integrity during critical missions. In addition, further studies are recommended to integrate other physiological metrics, such as electrocardiography (ECG) for heart rate variability analysis and electroencephalography (EEG) for direct monitoring of cortical activity, with the goal of creating multimodal monitoring systems that are even more precise.

Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

Abadi, R. V. and E. Gowen (2004). Characteristics of saccadic intrusions. *Vision Research 44*, 2675–2690.

Babu, M. D., D. V. JeevithaShree, G. Prabhakar, K. P. S. Saluja, A. Pashilkar, and P. Biswas (2019). Estimating pilots' cognitive load from ocular parameters through simulation and in-flight studies. *Journal of Eye Movement Research 12*(3), 1–16.

Biswas, P. and G. Prabhakar (2018). Detecting drivers' cognitive load from saccadic intrusion. *Transportation Research Part F: Traffic Psychology and Behaviour 54*, 63–78.

Caldwell, J. A. (2003). Fatigue in aviation operations. Technical report, Air Force Research Laboratory. Retrieved August 9, 2005.

Gawron, V. J. (1998, June). Human factors issues in the development, evaluation, and operation of uninhabited aerial vehicles. In *Proceedings of the AUVSI '98*, Huntsville, AL. Association of Unmanned Vehicle Systems International. June 8–12, 1998.

Hafed, Z. M. and J. J. Clark (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Research 42*, 2533–2545.

Hart, S. G. and L. E. Staveland (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human Mental Workload*, pp. 139–183. Amsterdam: North-Holland.

Landry, S. J. (Ed.) (2018). *Handbook of Human Factors in Air Transportation Systems*. Human Factors and Ergonomics. Boca Raton, FL: CRC Press. Taylor & Francis Group.

Liu, H., H. Wang, H. Jin, M. Ma, T. Pan, Z. Hu, M. Chu, and H. Si (2022). An improved algorithm for saccadic intrusion qualification under air traffic control situation. *International Journal of Industrial Ergonomics 92*, 103378.

McCarley, J. S. and C. D. Wickens (2005). Human factors concerns in uav flight. Technical report, Federal Aviation Administration. Retrieved July 30, 2005.

Sarmento, A. G. P., T. R. de Paula, A. S. de Oliveira, E. T. da Silva, J. Possamai, H. C. Marques, M. M. Cardoso Junior, and E. Villani (2022). A human-machine interface analysis for teleoperation of uav overtime delay. *Proceedings of the International Council of the Aeronautical Sciences (ICAS)*.

Shaffer, D. M., C. M. Krisky, and J. A. Sweeney (2003). Frequency and metrics of square-wave jerks: Influences of task-demand characteristics. *Investigative Ophthalmology & Visual Science 44*(3), 1081–1087.

Tattersall, A. J. and P. S. Foord (1996). An experimental evaluation of a subjective workload measuring technique. *Ergonomics 39*(6), 740–753.

Tokuda, S., G. Obinata, E. Palmer, and A. Chaparro (2011, August). Estimation of mental workload using saccadic eye movements in a free-viewing task. In *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, Boston, MA, USA. August 30–September 3.

Tokuda, S., E. Palmer, E. Merkle, and A. Chaparro (2009). Using saccadic intrusions to quantify mental workload. In *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*, Santa Monica, CA. Human Factors and Ergonomics Society.

Vidulich, M. A. and P. S. Tsang (1987). Absolute magnitude estimation and relative judgments in workload assessment. *Proceedings of the Human Factors Society 31st Annual Meeting*, 1506–1510.

Wickens, C. D. and J. G. Hollands (2000). *Engineering Psychology and Human Performance* (3 ed.). Upper Saddle River, NJ: Prentice Hall.