Materials Research Proceedings 37 (2023) 754-757 https://doi.org/10.21741/9781644902813-161

A brief review of pilots' workload assessment using flight simulators: subjective and objective metrics

Giuseppe Iacolino[1, a *], Antonio Esposito[1, b], Calogero Orlando[1, c] and Andrea Alaimo[1, d]

[1]Faculty of Engineering and Architecture, Kore University of Enna, 94100 Enna, EN, Italy

[a]giuseppe.iacolino@unikorestudent.it, [b]antonio.esposito@unikore.it, [c]calogero.orlando@unikore.it, [d]andrea.alaimo@unikore.it

Keywords: Human Factor, Subjective Methods, Objective Methods, Full Flight Simulator

Abstract. This study explores subjective and objective metrics to assess pilots WL, with a particular focus on the use of full flight simulators (FFS). The results show that FFS-based research demonstrates no significant differences compared to real flight experiences, highlighting the validity of FFS as a tool for studying pilots' performance. However, further research is needed to understand the impact of other parameters on pilots' performance, and to address human factor-related risks for enhanced aviation safety.

Introduction

Nowadays, workload (WL) can be considered one of the biggest challenges in the aeronautical field. The continuous growth of air traffic has led pilots to work under pressure and for extended periods of time. As reported in [1], 75% of aircraft accidents are linked to human errors and most of them are related to high levels of mental WL and fatigue. This is important in critical situations such as take-off, landing or emergency procedures. WL is associated with the onset of stress, a phenomenon with multiple facets ranging from neuroendocrine to psychological [2]. Pilots require training to effectively respond to both internal and external stimuli. Hence, using a flight simulator represents the most reliable, safe, time and cost savings way to study pilot's performances today [3]. Consequentially, it is necessary to employ some metrics that can better capture the pilots' stress level while performing the task. To do this, subjective and objective methods are widely exploited, and this work aims to describe such measures, with particular attention paid to the studies conducted using a full flight simulator (FFS).

Subjective and Objective Metrics

The concept of mental WL is described in [4] as a collection of mental, and composite brain states that influence human performance in various perceptual, cognitive, and sensorimotor abilities. Thus, it is crucial to choose the right metrics to evaluate it. Subjective and objective measures represent an important contribution on pilots' WL level evaluation. In detail, subjective WL measures comprise some tests which are administered while, before or after the performance. Being such subjective, they can be highly influenced by the pilot's psychological condition. That's why it is good practice to link them with the objective ones which represent the physical condition of the subject. The objective measures are mainly obtained via sensors which can detect some physiological parameters. In this work, subjective methods will be described first and objective ones after. The most reliable subjective methods are here mentioned.

The *NASA Task Load Index (TLX)* was first developed by Hart and Staveland in 1988. The TLX is widely recognized and extensively employed globally. The test has three stages to evaluate the pilot's WL. The first one obtains a global index across six dimensions: mental, physical, time demands, performance, effort, and frustration. Each dimension is divided into twenty intervals. In the second stage, pairs of dimensions are assessed, determining the ones with the greatest impact on WL. The third stage assigns scores to selected dimensions, with 0 for insignificant dimensions and 5 for the most important one [5]. On the other hand, the *Subjective Workload Assessment Technique (SWAT)* involves obtaining subjective ratings from participants regarding three dimensions of WL: time load, mental effort load, and psychological stress load. During the activity, participants are encouraged to assess their workload across three dimensions and rate it on a scale of low, medium, or high. A mental workload range scale is then calculated by combining the scores from all three dimensions. This assessment method relies on the self-report of participants [6]. Other subjective techniques are not mentioned for the sake of conciseness.

Regarding the objective metrics, some of them are reported here as well. As evidence, in [7] the eye movement provides indications on the pilot's visual perception, or [8] aimed to investigate the potential utility of α −Amy levels as a biomarker for stress in pilots operating within a high-stress environment. However, the heart rate variability (HRV) is one of the most employed parameters today. HRV is the spontaneous fluctuation in time between consecutive heartbeats, as measured by the distance between two successive R peaks on an electrocardiogram. This parameter can be studied in both frequency and time domain. In particular, the time domain measures also include the NN interval series standard deviation (SDNN). Based on the existing literature, a reduction in this parameter indicates an elevation in both mental WL and physical demands [9]. As previously mentioned, the study of HRV extends to the frequency domain. In this context, the fast Fourier transform is employed to estimate the Power Spectral Density associated with frequency bands, specifically focusing on high frequency and low frequency. These frequency bands provide valuable information about sympathetic and parasympathetic activities [10].

FFS for Human Factor Studies

Due to the rarity of FFS, there is a limited body of literature that specifically addresses their utilization. Most studies, in fact, primarily focus on simulators that provide visual cues to pilots through screens but lack the comprehensive flight sensation that an FFS can deliver. In 2018, the European Union Aviation Safety Agency (EASA) released the latest specifications for Airplane Flight Simulation Training devices (CSFSTD-A), which classify flight simulators into four levels of qualification: A, B, C, and D. The D-level FFS are the most advanced and reliable. For instance, to achieve a D-level qualification, the FFS must include a real-time feedback tool that allows the instructor to monitor the training envelope and prevent the airplane's operating limits from being exceeded. In [11] it is reported that acute effects on HRV and anxiety during a real-time flight are not significantly different from those experienced in a simulator, indicating that the simulated task planning and design closely approximate real-world conditions; in this case, an operational F-5M by Indra Company flight simulator was used. It is also reported that future flight simulators should incorporate immersive virtual reality technology simulating G forces and vibration. Nevertheless, in today's context, D-level FFS can be used to do reliable research and prevent in real-life risky decisions. In essence, pilots must be conscious on which kind of decisions they have to take while flying, both in high and normal WL condition. Additionally, the study conducted in [12] employed the CESSNA Citation C560 XLS FFS. The main findings of the study demonstrated that WL levels can be effectively differentiated by analyzing various domains of HRV, including time, frequency, and non-linear measures. These results highlight the importance of HRV indexes in assessing WL, and suggest the potential development of real-time, non-invasive instruments for evaluating it. In a separate study [13], the application of an FFS was investigated within typical flight scenarios. The research collected pilots' objective and subjective metrics to establish an evaluation index system. Research has revealed that pilot's errors have become a significant obstacle, impeding the progress towards enhancing flight safety within the aviation industry. The study also found some deficiencies in the design layout of the cockpit as in the flight crew operation process that cause pilots to make mistakes easily. The authors proposed that *designing aircrafts suitable* for pilots instead of *aircraft requiring pilots to adapt* can make the civil aircraft cockpit more humanized and highly automated. Also, the authors proposed a feasible method to analyze human factors (HF) in typical faults and incidents of transport category airplanes using comprehensive methods that combine subjective and objective measurements. In [14] it is investigated the correlation between subjective and objective indicators of fatigue, factors such as WL and work scheduling. It also examined whether the WL experienced by pilots during a simulator mission could serve as a moderator for increased fatigue after the mission. The study utilized a JAR STD 1A FFS. Results indicated that both subjective and objective measures of fatigue significantly rose during the three-hour experimental procedure. These findings suggest the presence of enduring effects from sleep deficit and propose a multifactorial model for assessing fatigue risks. In [15] two subjective WL measurements and three psychophysiological measurements were compared in both a simulator and a flight test. The comparisons were made across three flight scenarios using an ARJ21-700 FFS and a corresponding aircraft. Both flight scenarios and the flight environment significantly influenced NASA-TLX, eye blink rate, and HRV. Moreover, strong correlations were observed between the NASA-TLX and HRV, between the simulator and the flight test. These findings suggest that NASA-TLX and HRV can serve as consistent measures of WL in both FFS and real flight tests.

Conclusions

This short review aimed to highlight some subjective and objective metrics in order to understand the impact of stress on pilots' performance. The findings indicate that subjective and objective metrics play an important role for better understanding pilots' WL level. Furthermore, research involving FFS has shown promising results and benefits when compared to real flight experiences. Additionally, to emphasize the need for more studies is important. Despite current high safety standards, incidents related to HF still occur, highlighting the importance of ongoing investigation and improvement. By conducting more research and continuously improving understanding, pilot performance can be enhanced, risks can be mitigated, and safer aviation operations can be ensured for everyone involved.

Acknowledgment

This work was partially funded by the project "SAMOTHRACE" (Sicilian MicronanoTech Research And Innovation Center) – Funds: MUR, PNRR-M4C2, ECS_00000022, activity of Spoke 3 affiliated partner Kore University of Enna: WP3.4 SAfety MObility Sensors TEsting Platform - SA.MO.S.TE.P.

References

[1] Kharoufah, H., Murray, J., Baxter, G., & Wild, G. (2018). A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016. *Progress in Aerospace Sciences*, *99*, 1-13. https://doi.org/10.1016/j.paerosci.2018.03.002

[2] Masi, G., Amprimo, G., Ferraris, C., & Priano, L. (2023). Stress and Workload Assessment in Aviation—A Narrative Review. *Sensors*, *23*(7), 3556. https://doi.org/10.3390/s23073556

[3] Mohanavelu, K., Poonguzhali, S., Ravi, D., Singh, P. K., Mahajabin, M., Ramachandran, K., & Jayaraman, S. (2020). Cognitive Workload Analysis of Fighter Aircraft Pilots in Flight Simulator Environment. *Defence Science Journal*, *70*(2). https://doi.org/10.14429/dsj.70.14539

[4] Kramer, A. F., & Parasuraman, R. (2007). Neuroergonomics: Applications of neuroscience to human factors.

[5] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

[6] Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In *Advances in psychology* (Vol. 52, pp. 185-218). North-Holland. https://doi.org/10.1016/S0166-4115(08)62387-0

[7] Oberhauser, M., & Dreyer, D. (2017). A virtual reality flight simulator for human factors engineering. *Cognition, Technology & Work*, *19*, 263-277. https://doi.org/10.1007/s10111-017-0421-7

[8] Iizuka, N., Awano, S., & Ansai, T. (2012). Salivary alpha-amylase activity and stress in Japan air self-defense force cargo pilots involved in Iraq reconstruction. *American Journal of Human Biology*, *24*(4), 468-472. https://doi.org/10.1002/ajhb.22247

[9] Watson, D. W. (2001, June). Physiological correlates of heart rate variability (HRV) and the subjective assessment of workload and fatigue in-flight crew: a practical study. In *2001 People in Control. The Second International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres* (pp. 159-163). IET. https://doi.org/10.1049/cp:20010453

[10] Bilan, A., Witczak, A., Palusiński, R., Myśliński, W., & Hanzlik, J. (2005). Circadian rhythm of spectral indices of heart rate variability in healthy subjects. *Journal of electrocardiology*, *38*(3), 239-243. https://doi.org/10.1016/j.jelectrocard.2005.01.012

[11] Fuentes-García, J. P., Clemente-Suárez, V. J., Marazuela-Martínez, M. Á., Tornero-Aguilera, J. F., & Villafaina, S. (2021). Impact of real and simulated flights on psychophysiological response of military pilots. *International Journal of Environmental Research and Public Health*, *18*(2), 787. https://doi.org/10.3390/ijerph18020787

[12] Alaimo, A., Esposito, A., Faraci, P., Orlando, C., & Valenti, G. D. (2022). Human Heart-Related Indexes Behavior Study for Aircraft Pilots Allowable Workload Level Assessment. *IEEE Access*, *10*, 16088-16100. https://doi.org/10.1109/ACCESS.2022.3145043

[13] Xing, G., Sun, Y., He, F., Wei, P., Wu, S., Ren, H., & Chen, Z. (2023). Analysis of Human Factors in Typical Accident Tests of Certain Type Flight Simulator. *Sustainability*, *15*(3), 2791. https://doi.org/10.3390/su15032791

[14] Hörmann, H. J., Gontar, P., & Haslbeck, A. (2015). Effects of workload on measures of sustained attention during a flight simulator night mission.

[15] Zheng, Y., Lu, Y., Jie, Y., & Fu, S. (2019). Predicting workload experienced in a flight test by measuring workload in a flight simulator. *Aerospace medicine and human performance*, *90*(7), 618-623. https://doi.org/10.3357/AMHP.5350.2019