Training, Education, and Learning Sciences, Vol. 109, 2023, 202–210
https://doi.org/10.54941/ahfe1003166

Performance vs. Workload Matrix of Primary Flight Training: Exploratory Study

Gajapriya Tamilselvan, Nurettin Dinler, and Stephen M. Belt

Aviation Science, Saint Louis University, Saint Louis, MO 63103, USA

ABSTRACT

Objective: To explore the relationships between the performance and workload of pilots in a primary flight training environment.
Method: In this exploratory study, we measured physical workload parameters by recording the flight control deflections of the elevator and aileron and how they deviated from reference pitch and bank attitudes. We quantified flight performance by computing deviations between actual and desired altitude/heading parameters. Our study included a sample of twenty students and flight instructors from a Part 141 flight training school. Experimental stimuli for participants involved three instrument flight sessions in an Advanced Aviation Training Device (AATD) with ceiling and visibility set to unlimited, calm winds and light turbulent conditions. Participants were briefed to complete a preset flight pattern with seven segments twice per session. Flight segments were straight-and-level flight, level turns at a rate of 3 deg/s, a 500 ft/min climb and descent, and two airspeed changes. Spearman's correlation tests were used to examine the relationships between performance and workload data between sessions.
Results: Significant relationships between flight performance and physical workload parameters emerged from the data. Elevator workload was positively correlated with altitude performance across all flight sessions. There were positive relationships between elevator workload and heading performance during the first two sessions and no significant relationship in the last session. The aileron workload was inversely related to how much the pilots deviated from desired altitude performance. Aileron workload and heading performance were inversely related during the first and the last sessions, except for the second session. The research findings were limited in relation to generalizability to the population.
Conclusion: This study's results provide deeper insights into how pilots' performance relates to physical workload parameters in a primary flight training setting. This study's information elucidates the flight training community about skill development among Part 141 pilots and further provides a framework to develop evidence-based training strategies. Future research focuses on classifying the pilots' performance and workload into high/medium/low categories, investigating the nature of relationships, developing interactions, and relating them to pilot demographics.

Keywords: Flight training, Part 141, Workload, Performance, AATD

INTRODUCTION

The workload is a well-established metric to evaluate how pilots meet their task demands in flight and is intertwined with human performance. Flying an airplane is a complex task that places a demand on several aspects of a pilot's cognitive capabilities (Wilson and Hankins, 1994; Wilson, 2002; Causse et al., 2015). The impact of pilots' workload on performance and flight safety served as the subject of much systematic investigation (Škvareková, Pecho and Fedáš, 2021). Whilst some researchers identified a clear and negative association between pilot performance and workload, others reported dissociation and/or nonlinear correlation. Svensson and Wilson (2002) indicated significant relationships between heart rate and workload ratings, mental capacity, situational awareness, and performance. They found that mission job complexity increased workload, which in turn affected situational awareness (SA) and pilot performance.

Marris and Leung (2006) found that pilots experienced increasing difficulty in performing their flight tasks, as the workload increased. They reported that medium and high mental workload conditions had a negative influence on pilots' ability to listen, comprehend, and respond to auditory instructions. Borghini et al. (2014) conducted a comprehensive literature review of studies related to neurophysiological measurements (e.g., EEG, EOG, and heart rate) in pilots/drivers during their driving tasks. Neurophysiologic variables were correlated to the mental states of car drivers or airplane pilots during their control of the vehicles. The high mental workload was associated with the increased EEG power in the theta band and the reduced EEG power in the alpha band (Borghini et al., 2014).

Marinescu et al. (2016) investigated the relationship between mental workload, performance variance, and physiological measures by collecting data from various physiological measurements and subjective ratings of workload (e.g., the instantaneous self-assessment workload scale (ISA). Pilot performance measured within the task appeared to be negatively correlated with ISA ratings, indicating that as the mental workload increased, performance decreased (Marinescu et al., 2016). Marinescu et al. (2018) examined the relationship between experienced mental workload and physiological response by noninvasive monitoring of physiological parameters. Data for this study were collected using physiological measurements (heart interbeat intervals, breathing rate, pupil diameter, facial thermography), subjective ratings of workload (instantaneous self-assessment workload scale [ISA] and NASA-task load Index), and performance (Marinescu et al., 2018). Likewise, evidence was found for a negative association between performance and subjective workload, suggesting that respondents' task performance decreased as their subjective level of mental workload increased (Marinescu et al., 2018). A recent work examined the relationship between pilot workload, performance, subjective fatigue, sleep duration, number of sectors, and flight duration during short-haul operations and indicated weak, but significant correlations between workload and all factors (Arsintescu et al., 2020).

Pilots reported higher workloads as fatigue levels grew, the number of sectors increased, and objective performance worsened (Arsintescu et al., 2020). Hebbar et al. (2021) attempted to estimate cognitive workload using pilots' physiological indications (e.g., electroencephalographic (EEG) signals, ocular parameters, and pilot performance-based quantitative metrics). Introducing a secondary task along with flying led to a considerable increase in pilots' cognitive workload, resulting in a decrease in performance (Hebbar et al., 2021). Mansikka, Virtanen and Harris (2019) argued that the association between pilot performance and mental workload is more complex. Svensson et al. (1997) reported a non-linear relationship between pilot mental workload and performance, suggesting that mental workload influenced multiple components of pilot performance. In a study investigating the aircraft pilots' performance, mental workload, and tactical task goal awareness in a virtual flight training device, Mansikka, Virtanen and Harris (2019) found that when the pilot's awareness of the tactical goals was low, a combination of low performance and low mental workload occurred. Alaimo et al. (2020) found a complex and nonlinear relationship between workload, biometric data, and performance. It was impossible to assess the pilots' workload levels using merely subjective measurements in the context of aviation (Alaimo et al., 2020).

Nicholson et al. (1970) found a significant difference between the workload of the pilots during the let-down, approach, and landing. Hart and Hauser (1987) examined pilot workload using three workload measurements - communications performance, subjective ratings, and heart rate- to ascertain differences in flight-related task demands across various flight segments. Pilot ratings of workload, stress, and efforts appeared to be highly correlated and varied across flight segments, peaking during takeoff and landing (Har & Hauser 1987). Wilson and Hankins (1994) investigated the levels of pilots' cognitive workload during VFR and IFR flights using subjective and electroencephalographic (EEG) measures and reported that IFR flight segments were associated with higher levels of activity than VFR flight segments. Hankins and Wilson (1998) reported that pilots' heart rates increased during takeoffs and landings and to an intermediate level during instrument flying rules (IFR) segments and that pilots' brain waves indicated increased power during those flight segments, suggesting a higher workload for the pilots. Di Nocera, Camilli and Terenzi (2007) examined the professional pilots' eye movements during the different phases of a simulated flight (departure to landing) using spatial statistics algorithms. The results indicated that spatial dispersion indices were more sensitive to changes in mental workload during departure and landing, less sensitive during climb and descend, and least sensitive during the cruise phase (Di Nocera, Camilli and Terenzi, 2007).

Causse et al. (2012) explored the pilot's mental workload and psychological stress and their relationships with piloting activity and heart rate during the various flight segments (e.g., take-off, climb, cruise, approach, and landing). They found higher mental workload and stress levels for take-off and landing in comparison to other flight segments as well as a significant positive correlation between heart rate and mental workload/stress levels (Causse et al., 2012). Harbour et al. (2013) explored pilot workload and situation awareness in-flight during various phases of airborne operation on a tactical airlift aircraft and found significant changes in heart activity across flight segments. They identified assault landings, airdrops, and instrument approaches flown without a head-up display (HUD) as primary areas of concern for pilot stress and workload (Harbour et al., 2013). Alaimo et al. (2018) attempted to estimate the pilot's workload during two different flight segments (i.e., take-off-climb and approach-landing), to ascertain the feasibility of using low-cost noninvasive biometric devices as a sensor of pilot mental demand. Results showed that the workload level during the approach and landing phase was higher than the take-off and climb (Alaimo et al., 2018).

Agha (2020) evaluated the pilot workload during unexpected flight conditions (i.e., startle-thunder sound and surprise-engine failure). Single- and multi-engine aircraft were flown in a scenario that caused an uninformed surprise emergency condition, an uninformed surprise and startle emergency condition, and an informed emergency situation. Pilots' heart and respiration rates, flight performance, and subjective workload measures were collected during each condition. For both aircraft, the startle and surprise situations resulted in the highest heart and respiration rates, indicating an increased pilot workload. Furthermore, under all situations, the subjective assessments of mental, physical, and temporal workload, effort, and frustration were higher for twin-engine aircraft than for single-engine aircraft. The startle and surprise were measured using physiological indicators such as heart rate and respiration rate (Agha, 2020). Škvareková, Pecho and Fedáš (2021) sought to measure pilot workload using the heart rate variability (HRV) parameter during precision (Instrument Landing System-ILS) and non-precision (Non-Directional Beacon-NDB) approaches. The analysts reported that there was a difference between the workload of the pilots during different approaches (Škvareková, Pecho and Fedáš, 2021). More particularly, it was found that, except for one subject, pilots were under a higher workload during the ILS approach than the NDB approach (Škvareková, Pecho and Fedáš, 2021). In summary, there was a negative association between pilot workload and performance, and that pilot workload varies during different flight phases from low to high. However, the link between pilot effort and performance also appears to be complex.

Psychophysiological data from electroencephalogram (EEG) and flight simulator performance data were correlated to explore the relationships between the mental and physical workload of pilots (Belt et al., 2021). Existing literature did not classify whether the pilots are overworked or underworked and how that relates to their inflight task demands and experience. The purpose of this study was to explore the relationships between the performance and workload of pilots in a primary flight training environment.

METHODS

Participants

Our study included a sample of twenty students and flight instructors from a Part 141 flight training school: 2 student pilots, 11 private pilots and 7 commercial pilots. The mean of flight hours reported by the participants was:

239.18 total time, 191.9 pilot-in-command (PIC) hours, 4.95 actual instrument hours, 32.79 simulated instrument hours, and 31.44 flight training device (FTD) hours. Half of the sample comprising all commercial and three private pilots was instrument-rated prior to our experiment. Participants were required to be at least 18 years old to take part in this study.

Materials and Stimuli

Participants completed three simulated flight sessions on a Precision Flight Controls modular Flight Deck (PFCMFD) enabled with X-Plane 9 software. They were briefed to complete a preset flight pattern with seven segments twice per session (Figure 1). Flight segments were straight-and-level flight, level turns at a rate of 3 deg/s, a 500 ft/min climb and descent, and two airspeed changes. We extracted data from the software related to physical workload and flight performance. We measured physical workload parameters by recording the flight control deflections of the elevator and aileron and

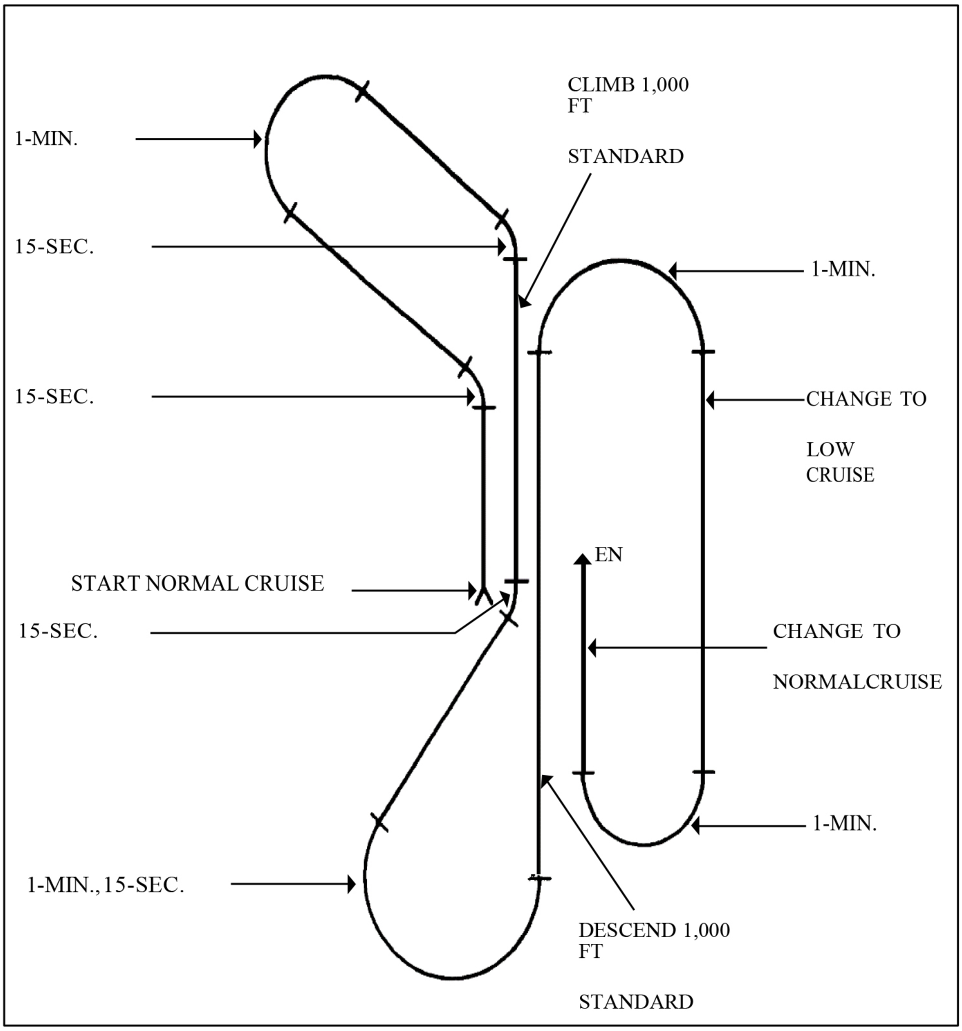

Figure 1: Flight pattern.

how they deviated from reference pitch and bank attitudes. We quantified flight performance by computing altitude and heading performance and how they deviated between the actual and desired parameters. Spearman's correlation tests were used to examine the relationships between performance and workload data between three sessions.

Design and Ethics

This study employed a quantitative correlational design to investigate how performance and workload were related in a primary flight training setting. This research was approved by the Institutional Review Board (IRB# 24183) to fulfill legal and ethical considerations.

RESULTS & DISCUSSION

Spearman's correlation resulted in significant relationships between performance and workload in a primary flight training setting. As our hypotheses were non-directional, we selected two-tailed tests for conducting non-parametric correlation tests. Spearman's Rho correlation coefficients for all three trials were listed in Tables (1-3).

The elevator workload positively relates to and accounts for a 2.19% variance in altitude performance. Elevator workload positively relates to and accounts for 0.16% variance in heading performance. Aileron workload inversely relates to and accounts for a 3.02% variance in altitude performance. Aileron workload inversely relates to and accounts for 0.18% variance in heading performance.

The elevator workload positively relates to and accounts for a 2.02% variance in altitude performance. Elevator workload positively relates to and accounts for 0.008% variance in heading performance. Aileron workload

Table 1. Performance vs. workload matrix for trial 1.

	Altitude Performance	Heading Performance
Elevator Workload	$r_s = .148, p < .001$	$r_s = .040, p < .001$
Aileron Workload	$r_s = -.174, p < .001$	$r_s = -.043, p < .001$

Table 2. Performance vs. workload matrix for trial 2.

	Altitude Performance	Heading Performance
Elevator Workload	$r_s = .142, p < .001$	$r_s = .009, p < .05$
Aileron Workload	$r_s = -.149, p < .001$	$r_s = .023, p < .001$

Table 3. Performance vs. workload matrix for trial 3.

	Altitude Performance	Heading Performance
Elevator Workload	$r_s = .253, p < .001$	$r_s = .006, p = .202$
Aileron Workload	$r_s = -.159, p < .001$	$r_s = -.011, p < .05$

inversely relates to and accounts for a 2.22% variance in altitude performance. Aileron workload positively relates to and accounts for 0.05% variance in heading performance.

Elevator workload positively relates to and accounts for a 6.4% variance in altitude performance. There was no significant relationship between elevator workload and heading performance. Aileron workload inversely relates to and accounts for a 2.53% variance in altitude performance. Aileron workload inversely relates to and accounts for 0.01% variance in heading performance.

Two prominent trends from the data analysis were: 1) heading and altitude performance increased with the increase in elevator workload, and 2) heading and altitude performance decreased with the increase in aileron workload. The second trend was consistent with the studies in the literature, where the workload was measured by subjective scales (Marinescu et al., 2018) and physiological sensors (Wilson, 2002; Hebbar et al., 2021). The current study quantified workload by measuring the deflections of the elevator and aileron to reduce any interventions.

CONCLUSION

The present study sought to investigate how flight performance data correlated with physical workload parameters experienced by pilots. Altitude performance rose with the increase in elevator workload across all flight sessions. But pilots' desired altitude performance declined when they exerted a higher aileron workload. Heading performance was positively correlated with elevator workload for the first two sessions. Pilots exhibited lower heading performance and higher aileron workload during the first and last sessions apart from the second session. These findings were representative of the study's sample and could not be generalized to a broader population. This study utilized a convenient sampling strategy to recruit participants, which further limited the findings by selection bias and sampling error. Quantifying the relationships between flight performance and workload parameters inform the flight training community about the skill development among Part 141 pilots. These efforts will provide a framework to develop evidence-based training strategies. Future research will stratify flight performance and workload into high/medium/low categories, examine interrelationships, and establish interactions with pilot demographics.

REFERENCES

Agha, R. D. (2020) *Evaluating scenarios that can startle and surprise pilots*. PhD Dissertation. Embry-Riddle Aeronautical University. Available at: https://commons.erau.edu/edt/510/.

Alaimo, A. *et al.* (2018) "A Pilot Mental Workload Case Study in a Full Flight Simulator," *Aerotecnica Missili & Spazio* [Preprint]. Available at : https://doi.org/10.1007/bf03404762.

Alaimo, A. *et al.* (2020) "Aircraft Pilots Workload Analysis: Heart Rate Variability Objective Measures and NASA-Task Load Index Subjective Evaluation," *Aerospace*, 7(9), p. 137. Available at : https://doi.org/10.3390/aerospace7090137.

Arsintescu, L. *et al.* (2020) "The relationship between workload, performance and fatigue in a short-haul airline," *Chronobiology International*, 37(9–10), pp. 1492–1494. Available at: https://doi.org/10.1080/07420528.2020.1804924.

Belt, S. *et al.* (2021) "Exploring Pilot Workload During Professional Pilot Primary Training and Development: A Feasibility Study," *Advances in Human Factors and System Interactions: Proceedings of the AHFE 2021 Virtual Conference on Human Factors and Systems Interaction, July 25-29, 2021, USA*, pp. 193–201. Available at: https://doi.org/10.1007/978-3-030-79816-1_24.

Borghini, G. *et al.* (2014) "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neuroscience & Biobehavioral Reviews*, 44, pp. 58–75. Available at: https://doi.org/10.1016/j.neubiorev.2012.10.003.

Causse, M. *et al.* (2012) "An analysis of mental workload and psychological stress in pilots during actual flight using heart rate and subjective measurements," *5th International Conference on Research in Air Transportation (ICRAT)* [Preprint]. Available at: https://oatao.univ-toulouse.fr/16202/.

Causse, M. *et al.* (2015) "EEG/ERP as a Measure of Mental Workload in a Simple Piloting Task," *Procedia Manufacturing*, 3, pp. 5230–5236. Available at: https://doi.org/10.1016/j.promfg.2015.07.594.

Di Nocera, F., Camilli, M. and Terenzi, M. (2007) "A Random Glance at the Flight Deck: Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload," *Journal of Cognitive Engineering and Decision Making*, 1(3), pp. 271–285. Available at: https://doi.org/10.1518/155534307x255627.

Hankins, T. C. and Wilson, G. F. (1998) "A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight.," *Aviation, Space, and Environmental Medicine*, 69(4), pp. 360–7.

Harbour, S. D. *et al.* (2013) "An integrated neuroergonomic assessment of inflight pilot workload," *17th International Symposium on Aviation Psychology*, pp. 621–626. Available at: https://corescholar.libraries.wright.edu/isap_2013/11/.

Hart, S. G. and Hauser, J. R. (1987) "Inflight application of three pilot workload measurement techniques.," *Aviation, Space, and Environmental Medicine*, 58(5), pp. 402–10.

Hebbar, P. A. *et al.* (2021) "Correlation Between Physiological and Performance-Based Metrics to Estimate Pilots' Cognitive Workload," *Frontiers in Psychology*, 12. Available at: https://doi.org/10.3389/fpsyg.2021.555446.

Mansikka, H. P., Virtanen, K. and Harris, D. (2019) "Dissociation Between Mental Workload, Performance, and Task Awareness in Pilots of High-Performance Aircraft," *IEEE Transactions on Human-Machine Systems*, 49(1), pp. 1–9. Available at : https://doi.org/10.1109/thms.2018.2874186.

Marinescu, A. *et al.* (2016) "Exploring the Relationship between Mental Workload, Variation in Performance and Physiological Parameters," *IFAC-Papers Online*, 49(19), pp. 591–596. Available at : https://doi.org/10.1016/j.ifacol.2016.10.618.

Marinescu, A. *et al.* (2018) "Physiological Parameter Response to Variation of Mental Workload," *Human Factors*, 60(1), pp. 31–56. Available at: https://doi.org/10.1177/0018720817733101.

Nicholson, A. *et al.* (1970) "Activity of the nervous system during the let-down, approach and landing: A study of short duration high workload," *Clinical Aviation and Aerospace Medicine,* 41(4), p. 436–446. Available at: https://apps.dtic.mil/sti/citations/AD0719980.

Svensson, E. *et al.* (1997) “Information complexity-mental workload and performance in combat aircraft,” *Ergonomics*, 40(3), pp. 362–380. Available at: https://doi.org/10.1080/001401397188206.

Svensson, E. and Wilson, G. F. (2002) “Psychological and Psychophysiological Models of Pilot Performance for Systems Development and Mission Evaluation,” *The International Journal of Aviation Psychology*, 12(1), pp. 95–110. Available at: https://doi.org/10.1207/s15327108ijap1201_8.

vSkvareková, I., Pecho, P. and Fedáš, J. (2021) “Measurement of Pilot’s Workload During NDB and ILS Approaches,” *Transportation Research Procedia*, 59, pp. 14–20. Available at : https://doi.org/10.1016/j.trpro.2021.11.092.

Wilson, G. D. (2002) “An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures,” *The International Journal of Aviation Psychology*, 12(1), pp. 3–18. Available at: https://doi.org/10.1207/s15327108ijap1201_2.

Wilson, G. D. and Hankins, T. (1994) “Eeg and Subjective Measures of Private Pilot Workload,” *Proceedings of the Human Factors and Ergonomics Society... Annual Meeting*, 38(19), pp. 1322–1325. Available at: https://doi.org/10.1177/154193129403801916.