Advances in Human Factors of Transportation, Vol. 186, 2025, 20–32
https://doi.org/10.54941/ahfe1006490

Mayday, Mayday! – Is Heart Rate Variability a Suitable Objective Indicator to Detect Pilot's Increased Mental Workload in Emergency Situations?

Giulia Troyer, Maria Hagl, and Jörn Jakobi

Institute of Flight Guidance, German Aerospace Center (DLR), 38108 Braunschweig, Germany

ABSTRACT

Human performance (HP) is a crucial factor in aviation, especially in high-cognitive-demand situations such as emergencies. This study explores the relationship between mental workload (MWL) and heart rate variability (HRV) in professional pilots, focusing on how these variables are influenced by different flight phases, pilot roles, and the implementation of an intelligent pilot assistance system (IPAS). Using a low-fidelity flight simulator, eight male pilots faced emergency scenarios, and their MWL was evaluated through subjective ratings. HRV was measured using the RMSSD parameter. Contrary to previous findings, this study's results did not reveal a correlation between MWL and RMSSD. However, RMSSD values showed significant variation across flight phases, with the lowest RMSSD observed during critical decision-making processes in the FORDEC phase, rather than during the emergency phase. These results suggest that while RMSSD may not demonstrate a direct correlation with MWL, it remains a versatile and non-invasive method for monitoring physiological states. Implications for the use of HRV in real-life operations and the assessment of new assistance systems are discussed.

Keywords: Mental workload, Heart rate variability, Aviation

INTRODUCTION

Human Performance and Adaptive Automation in Aviation

Human performance (HP) is vital for flight safety and efficiency. Pilots handle complex systems, demanding significant cognitive engagement. The complexity and simultaneous task processing affect the necessary cognitive resources. A tragic example is Air France 447, where incorrect system information impaired situational awareness (SA), increased workload, and led to conflicting decisions that caused a fatal crash. The computerized system, meant to reduce mental workload (MWL) (de Wit and Moraes Cruz, 2019), made issue detection nearly impossible. Advancing artificial intelligence (AI) could reduce cockpit ambiguities, but human-AI collaboration remains challenging. Billings (1996) stresses that humans and systems must continuously share their intentions for clear and direct communication. By deploying intelligent assistants (IA) into the cockpit, AI can enhance SA and reduce workload (Kirwan, 2024). In emergencies, pilots may develop tunnel vision (de Wit and Moraes Cruz, 2019), impairing SA. When HP declines, automation should provide support; however, during low-workload periods, it should recede – a concept known as human-centered automation (Sheridan, 1995). This approach requires humans and systems to be aware of each other's state (Billings, 1996), enabling adaptive support and dynamic task distribution (Parasuraman et al., 1999). In this context, a key prerequisite is the valid measurement of the operator state. MWL is part of the generic concept of workload, which includes mental, physical, and emotional aspects (Gaillard, 1993). This study refers to MWL only, as the cognitive demands in the cockpit are the primary concern. Although there is no universal definition, this paper defines MWL as the *"portion of the operator's limited capacity actually required to perform a particular task"* (O'Donnell and Eggemeier, 1986, p. 42)

Mental Workload and Stress

HP is influenced by MWL and other cognitive, emotional, and physiological states that can interact with it. MWL and stress, although often viewed as synonyms, must be differentiated. Both focus on environmental demands and operator capacity but stem from different theories: workload theories view humans as information processors handling tasks with resources, while stress theories highlight threats, opportunities, and individual responses in the work environment (Gaillard, 1993).

Subjective Mental Workload Measurements

Given MWL's role in predicting HP and enabling adaptive automation, how can we assess pilots' MWL reliably and practically? MWL, unobservable directly (Casali and Wierwille, 1984), must be derived from measurable indicators: physiological, subjective, and performance-based measures (Miller, 2001). Although performance-based measures are easy to apply and primary task data is often relevant, they may not always suit MWL assessment (Boumann et al., 2023). To reveal MWL variances, subjective and physiological measures should ideally be combined (Sammito et al., 2024). Subjective measurements are the most practical (Miller, 2001), convenient, time-efficient, and flexible (Yeh and Wickens, 1988). A key single indicator scale is the Instantaneous Self-Assessment (ISA, see Tattersall and Foord, 1996). According to the authors, it is comparable to other MWL assessments and measures like Heart Rate Variability (HRV).

Physiological Mental Workload Measurements

Assigning a physiological measure to a cognitive state is challenging, as each can reflect multiple cognitive states (Boumann et al., 2023). HRV, a widely used, non-invasive MWL indicator, is best assessed via electrocardiography (ECG). HRV refers to *"the change in the time intervals between adjacent heartbeats"* (McCraty and Shaffer, 2015, p. 46) and reflects (para-) sympathetic activity, providing insights into the autonomic nervous system

(ANS) function (Lennartsson et al., 2016). The parasympathetic nervous system (PNS), the relaxed response system, lowers the heart rate (HR) by releasing acetylcholine, while the sympathetic nervous system (SNS), the quick response system, increases HR via adrenaline and noradrenaline (Gordan et al., 2015). These systems maintain autonomic balance (Pham et al., 2021). Generally, a low HRV signifies increased sympathetic activity, whereas a high HRV indicates elevated parasympathetic activity (Lennartsson et al., 2016). Extensive research has explored the link between HRV and cognitive states. Green et al. (2016) suggest that HRV can detect physiological markers associated with stress. Taelman et al. (2009) found HR and HRV changes during mental tasks, making them promising stress measures. Similarly, describe HRV as *"a valid measure of the psychological stress response"*. Reduced parasympathetic activity is linked to decreased HRV under occupational stress (Järvelin-Pasanen et al., 2018, The et al., 2020, Thielmann et al., 2021). Beyond stress, HRV also reflects states of a more cognitive nature. Gonzalez et al. (2023) found HRV suitable to measure physical and cognitive fatigue, while Pham et al. (2021) report links to cognitive fatigue and affective dysregulation. But how evident is the relationship between HRV and MWL? MWL impacts cardiovascular functions (Delliaux et al., 2019) and can specifically influence HRV (Nagasawa and Hagiwara, 2016). Typically, increased MWL is associated with decreased HRV (Delliaux et al., 2019) assuming stable physical WL (Meshkati, 1988, Myrtek et al., 1999, Tattersall and Hockey, 1995). Consequently, HRV can be considered a simple MWL measure if respiratory influences are excluded (Miyake, 2001). Nevertheless, HRV is multifaced, and alterations cannot be solely attributed to MWL variations. DiDomenico and Nussbaum (2011) stress the need for careful HRV interpretation, as it varies by task type. Additional influencing factors include biological (e.g., age, gender), diseases, physical fitness, sports activity, increased body weight, smoking, alcohol abuse, and external factors like noise. Common HRV assessment methods include frequency- and time-domain analysis (McCraty and Shaffer, 2015). Time-domain parameters assess variance in inter-beat intervals (IBI) through statistical measures (McCraty and Shaffer, 2015), while frequency-domain parameters reflect power distribution across frequency bands (Pham et al., 2021). Frequency-domain parameters are excluded given the wider use and comparability of time-domain parameters. Cinaz et al. (2013) and Fallahi et al. (2016) demonstrate this relationship for the time-domain parameter 'Root Mean Square of Successive RR interval Differences' (RMSSD).

Research Questions and Hypotheses

This study re-examines the correlation between MWL and HRV (see Tattersall and Foord, 1996) through time-domain parameters in an operational context with real pilots using a low-fidelity cockpit simulation. It is postulated that higher MWL (ISA ratings) is associated with reduced HRV (RMSSD values) and vice versa. The hypothesis is formulated as follows:

H1: There is a correlation between pilots' ISA ratings and RMSSD during flight involving emergency scenarios.

Research suggests that flight crews' perceived MWL can depend on system interfaces (Martins, 2016). Moreover, both IA usage and emergency situations likely affect pilots' MWL and, accordingly, their HRV. In this study, pilots operated in flight crews consisting of a pilot flying (PF) and a pilot monitoring (PM). PF and PM have significantly different tasks (Dehais et al., 2017), which could lead to role-dependent HRV differences. Lastly, mean HR varies by flight phase (Lee and Liu, 2003), so here, simulator flights are divided into four phases: (1) initial, (2) emergency, (3) solution (FORDEC)[1], and (4) break phase (see Procedure below). It is assumed that the phase may also influence HRV. What effects do the use of IA, the pilot's role (PF vs. PM), and the flight phase have on HRV? The hypotheses are as follows:

H2: RMSSD is lower during the emergency phase compared to the other flight phases.

H3: RMSSD is lower in the PF-role compared to the PM-role.

H4: RMMSD is higher on flights with an IA than on flights without one.

Comparisons regarding H3 and H4 are restricted to the FORDEC phase, as this is where the critical decision-making occurs. H2 and H4 analyses only consider the PF's RMSSD values.

METHOD

Participants

Eight professional male pilots, aged 25 to 56 (M = 35.00, SD = 10.41), participated in the study. All were native German speakers and held an air transport pilot license (A), commercial pilot license (A), or multi-crew pilot license (A) with an Airbus-type rating. They had over 1500 flight hours. Participation was voluntary, with compensation for travel, lodging, and study time.

Material

This section first describes the flight simulator used, then the material to assess HRV and subjective MWL.

Flight Simulator

This research, part of a larger study, was conducted at the Institute of Flight Guidance at DLR using the iSIM flight simulator, which features touch-input displays to easily adjust cockpit controls. For the study, the Intelligent Pilot Assistance System (IPAS), developed at DLR, was used to support safety-critical emergency situations. *"The IPAS is an AI-based system intended to assist pilots in decision making by identifying and assessing situations and generating options for action at the strategic flight management level"* (Würfel et al., 2023, p. 591). According to the authors, it can operate in two modes: Mission Support for normal operations and Option Support Mode for abnormal situations, such as finding alternate routes/airports. Only the Option Support Mode was used here.

[1]FORDEC is a method in aviation that helps pilots to make decisions and stands for *"Facts, Options, Risks and Benefits, Decision, Execution, Check"* Hörmann, H.-J. 'Training of aircrew decision making'. AGARD, 15.

HRV Measurement

Per simulation run, each participant (PF and PM) wore a Polar H10 (Polar Electro Oy, Kempele, Finland) chest strap. The HRV data was analyzed using the Kubios HRV Scientific software (Tarvainen et al., 2014). The Polar H10 straps were connected via Bluetooth to the Kubios HRV app on iPads.

Subjective MWL Measurement

Subjective MWL was assessed using the ISA scale, with participants stating the corresponding number aloud. MWL was measured twice within each flight (ISA 1: after emergency, ISA 2: after FORDEC), with four flights in total for altering pilot role and IPAS support (on/off), resulting in eight ISA-ratings per participant.

Procedure

Participants were briefed on the study purpose and HRV measurement. The Polar H10 strap was applied, and they were introduced to the iSIM, IPAS software, and the procedure as a two-person crew, starting with one as PF and the other as PM. The experiment involved two distinct flight scenarios[2], each with emergency situations not briefed to the pilots. Scenario 1 was scripted with an in-flight engine failure, while scenario 2 simulated a fuel leak. All scenarios started at cruising altitude, approximately 1–2 minutes before the emergency onset. Participants were instructed to operate as they would during a regular flight, which contained following the Electronic Centralized Aircraft Monitor (ECAM) actions in an emergency situation and initiating the FORDEC procedure after that. Each scenario was conducted twice- once with and once without the IPAS- while alternating pilot roles, resulting in four simulator flights in total. The order of flights was always: Flight A (scenario 1 / non-IPAS), Flight B (scenario 2 / non-IPAS), Flight C (scenario 1 / IPAS), and Flight D (scenario 2 / IPAS). Each flight took approximately 13 minutes. After each flight, there was a short 15-minute break following the first two flights. Each flight proceeded as follows: In the first 1–2 minutes, pilots familiarized themselves with their surroundings. Once the emergency occurred, participants initiated ECAM actions, while one experimenter acted as ATC. After completing the ECAM actions, the first ISA rating was requested (ISA 1). Participants then used FORDEC, employing IPAS as needed, and decided on an alternate airport before the second ISA rating (ISA 2). The scenario continued for another 30 seconds to 1 minute, with no landing required in this study.

Study Design and Data Analysis

The study follows a 2 (IPAS: with vs. without) × 2 (pilot role: PF vs. PM) × 4 (flight phase: initial vs. emergency vs. FORDEC vs. break) design with repeated measurement. The dependent variables are MWL (measured by the ISA-scale) and HRV (measured by RMSSD). All data analyses were performed using R (R Core Team, 2024, v4.4.0). To test

[2]The two scenarios are relatively similar in length and difficulty, which is why they are not expected to influence HRV and are therefore not included in the study design.

H1, a correlation with repeated measurement (rmcorr) was conducted using the rmcorr-package in R (Bakdash and Marusich, 2017). The rmcorr was calculated separately for ISA1 - RMSSD1 (rmcorr1) and for ISA2 - RMSSD2 (rmcorr2). A mixed repeated measures ANOVA was initially planned for the additional research questions. Since statistical assumptions were not met, the corresponding hypotheses were tested separately. For H2, a Friedman test was conducted, followed by pairwise post hoc t-tests with Bonferroni correction. For H3 and H4, pairwise t-tests were applied. Significance level was set at $\alpha = .05$.

RESULTS

H1: Correlation ISA – HRV

Across all conditions, the mean reported MWL was 2.55 (SD $ISA_{comb} = 0.75$). Descriptively, ISA 1 (after Emergency) was slightly lower than ISA 2 (after FORDEC) ($M_{ISA1} = 2.47$, $SD_{ISA1} = 0.92$ and $M_{ISA2} = 2.63$, $SD_{ISA2} = 0.55$). Results for rmcorr1 show no significant correlation between ISA1 and RMSSD1, r_{rm} (23) = .13, 95% CI [–.28, .50], $p = .53$. Similarly, the results for rmcorr2 show no significant correlation between ISA2 and RMSSD2, r_{rm} (23) = .29, 95% CI [–.11, .62], $p = .15$.

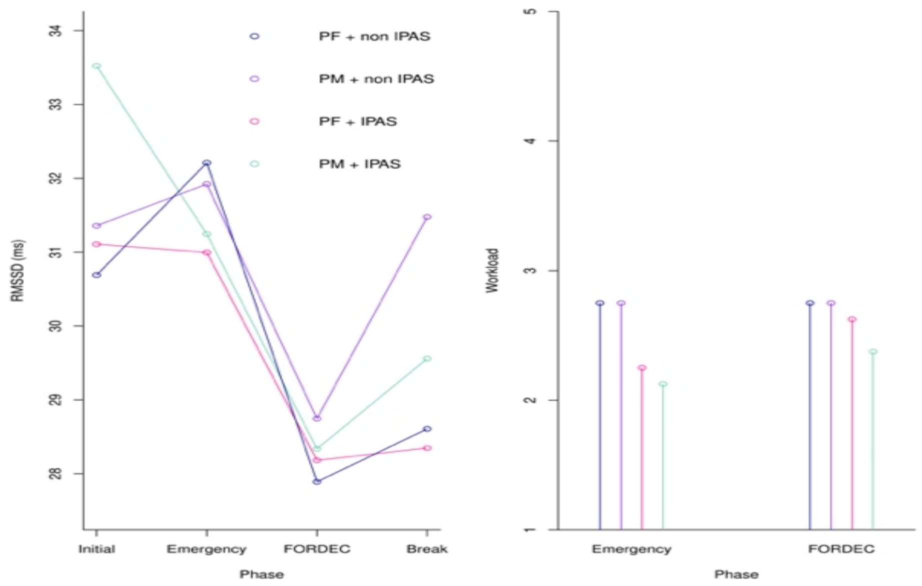

Figure 1: RMSSD and ISA ratings vary by flight phase, pilot role, and IPAS use. The right figure shows emergency and FORDEC phases where MWL was assessed.

H2-4: Impact of IPAS, Pilot Role, and Flight Phase on HRV

For H2, a significant difference was found in RMSSD values across the four phases in the simulation run, $\chi^2(3) = 26.42$, $p < .001$. Descriptive statistics show that RMSSD is lowest in FORDEC ($M = 28.26$; $SD = 9.76$), followed by break ($M = 28.95$; $SD = 8.83$), emergency ($M = 31.12$; $SD = 10.74$), and initial ($M = 32.32$; $SD = 10.14$). For post-hoc comparisons, the Bonferroni adjustment results in a level of significance at $\alpha_{adj} = .02$. As shown in **Table 1**, pairwise t-tests reveal that RMSSD is significantly lower in the FORDEC phase than in the emergency phase, $t(7) = 3.89$, $p < .01$; Cohen's $d = .28$, 95% CI [–.71 ; 1.26]. All other pairwise comparisons are non-significant.

Table 1: Results for the three pairwise comparisons with Bonferroni adjustment.

Group 1	Group 2	df	t	p
Emergency	Initial	7	1.59	.16
Emergency	FORDEC	7	3.89	<.01*
Emergency	Break	7	1.84	.10

Figure 1(left) descriptively shows RMSSD values according to the four phases and ISA ratings for each of the experimental conditions. As shown in **Figure 1(right)**, in all four conditions, the minimum RMSSD was reached in the FORDEC phase. Concerning H3, the average RMSSD values in the FORDEC phase were lower for the pilot flying ($M = 25.56$; $SD = 8.94$) than for the pilot monitoring ($M = 28.32$; $SD = 9.37$). This difference is not statistically significant $t(7) = -1.05$, $p = 33$.

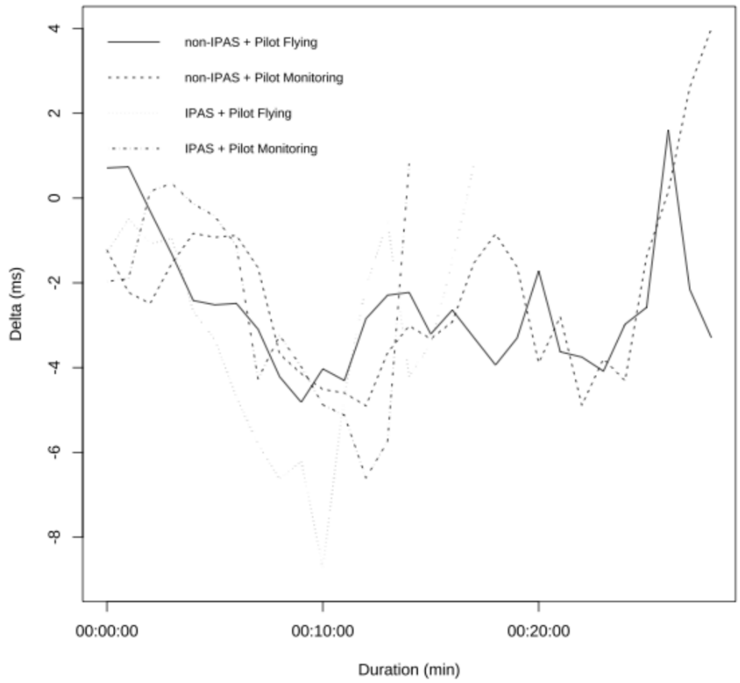

Figure 2: Visualization of HRV dynamics over time, depending on pilot role and IPAS use.

Regarding H4, the average RMSSD values for pilots flying during the simulation run in the FORDEC phase were lower when using the IPAS

($M = 28.18$; $SD = 9.72$) compared to when it was not used ($M = 28.34$; $SD = 10.04$). This difference is not statistically significant $t(7) = 0.14$, $p = .89$.

Exploration: HRV Dynamics Over Time

Changes in RMSSD over time are visualized in **Figure 2.** Notably, non-IPAS flights are considerably longer than non-IPAS flights. In the IPAS condition, RMSSD of the PF shows a notable decrease after the first few minutes. In contrast, RMSSD of the PM in the IPAS condition increases towards the end. For the non-IPAS condition, curves of the PF and the PM show a relatively similar pattern.

DISCUSSION

The present study involved pilots in an emergency scenario within a low-fidelity flight simulator to assess a possible correlation between pilots' MWL and HRV. No significant correlation was found. Furthermore, the study examined the effects of flight phase, pilot role and of an IA on HRV. Neither pilot role nor the use of IA significantly affected HRV. Pairwise t-tests reveal a significantly lower RMSSD in the FORDEC phase compared to the emergency phase.

Despite many studies showing a significant correlation between HRV and MWL (Delliaux et al., 2019, Green et al., 2016), no such correlation was found here. DiDomenico and Nussbaum (2008) indicate that HRV is less reliable at low workload but more sensitive with large muscle use in high workloads. In this study, pilots stayed seated and did not use large muscle groups. The overall low ISA ratings suggest a low workload, which is understandable considering the easily solvable emergency scenarios and the pilots' experience. A positive correlation between ISA and RMSSD contradicts previous findings (e.g., Tattersall and Hockey, 1995). As this correlation is non-significant, the positive valence should not be overemphasized.

Analyses of the additional research questions reveal significant differences in PF RMSSD values across all flight phases. Contrary to H2, RMSSD was not significantly lower during the emergency phase compared to the other three flight phases. Interestingly, the lowest RMSSD values are observed during the FORDEC phase. Although both emergency and FORDEC procedures are thoroughly practiced, emergencies tend to be sudden, likely elevating MWL levels more than FORDEC procedures do. In a flight simulator study with fighter pilots, Fernández-Morales et al. (2024) found low HRV in emergencies, which was interpreted as an anticipatory anxiety response of the nervous system. A possible explanation for our findings is that the simulated emergency may not have been perceived as a real threat. Additionally, pilots were instructed to focus on the FORDEC procedure. This requirement may have increased MWL during FORDEC, as pilots performed it more rigorously than usual. As lower HRV is generally associated with higher MWL (Delliaux et al., 2019), lower HRV in the FORDEC phase indicates higher MWL. This aligns with the high volume of information to be evaluated during the FORDEC phase (Li et al., 2011). For IPAS flights, using a new system could have been a crucial factor as pilots needed to adjust to the interface and alter their entire decision-making process. In fact, research shows that the interface of systems can have an impact on the perceived MWL of pilots (Martins, 2016). A last explanation for our findings could be the increased communication during FORDEC. External factors like ongoing communication noise and social stress (possibly related to the discussion about alternate airports) can decrease HRV (Sammito and Böckelmann, 2016, Schnell et al., 2013).

HRV dynamics over time show similar patterns between pilot roles for non-IPAS flights. Nevertheless, our results suggest that RMSSD in the non-IPAS condition is descriptively lower for PF than for PM. For IPAS flights, the same descriptive difference is observable. According to Hart and Hauser (1987), the overall HR on regular flights is higher for the PF than for the PM. This finding is consistent with our findings and makes sense, as the PF holds the primary responsibility. Furthermore, results indicate a considerably lower RMSSD for the PF in IPAS-flights. Such a decrease in RMSSD is not observed for the PM. The primary responsibility and new system usage likely increased cognitive load for the PF, associated with increased MWL. More training with new systems is needed to reduce novelty effects and enhance efficiency.

This study has various limitations. First, the fixed order of conditions likely influenced their performance, leading to faster emergency resolution. This limitation arises because this research was part of a larger study, which imposed constraints on the experimental design. It is possible that the HRV could have been even lower if they had experienced the emergencies for the first time when using the IPAS. Future studies should use a between-subjects design to avoid systematic training effects. Second, the sample size of eight participants was insufficient to detect small effects and to differentiate between covariates beyond the focused MWL. It would be desirable to repeat the study with a bigger sample size. Additionally, a higher variance in ISA would help to accurately correlate it with HRV. For that reason, the interpretation of the correlation is limited. Third, while HRV provides insights into sympathetic and parasympathetic activity, its temporal resolution is imperfect (Hoover et al., 2012). HRV can also indicate fatigue, stress, anxiety, burnout, and muscle activities (Immanuel et al., 2023). In future research, HRV measures should be combined with simple HR to obtain a more detailed picture (see Hirt et al., 2020).

CONCLUSION

Previous research indicates lower HRV correlates with increased MWL (see Delliaux et al., 2019), but this study found no correlation. While there are limitations, significant HRV variances were noted during critical flight phases such as FORDEC. The lack of correlation between HRV and subjective ISA complicates the interpretation of these variances in relation to MWL. Achieving this goal may be difficult, as various cognitive and physiological states can influence HRV. Thus, concentrating on HRV alone may be more effective. A significant decline in HRV could suggest a detrimental psychophysiological state if sustained. HRV measurements offer a flexible, non-invasive method for monitoring physiological states. They can also support developing new assistance systems and applying them in real-life operations, improving responsiveness to critical HP states and facilitating adaptive automation. Further research is required to validate this in predicting HP.

REFERENCES

Bakdash, J. Z. and Marusich, L. R. (2017) 'Repeated Measures Correlation', *Frontiers in Psychology,* 8. doi: https://doi.org/10.3389/fpsyg.2017.00456.

Billings, C. E. (1996) *Human-centered aviation automation: Principles and guidelines*.

Boumann, H., Hamann, A., Biella, M., Carstengerdes, N. and Sammito, S. 'Suitability of Physiological, Self-report and Behavioral Measures for Assessing Mental Workload in Pilots'. *Engineering Psychology and Cognitive Ergonomics*, Cham: Springer Nature Switzerland, 3–20.

Casali, J. G. and Wierwille, W. W. (1984) 'On the measurement of pilot perceptual workload: A comparison of assessment techniques addressing sensitivity and intrusion issues', *Ergonomics,* 27(10), pp. 1033–1050. doi: https://psycnet.apa.org/doi/10.1080/00140138408963584.

Cinaz, B., Arnrich, B., La Marca, R. and Tröster, G. (2013) 'Monitoring of mental workload levels during an everyday life office-work scenario', *Personal and Ubiquitous Computing,* 17(2), pp. 229–239. doi: https://doi.org/10.1007/s00779-011-0466-1.

de Wit, P. A. J. M. and Moraes Cruz, R. (2019) 'Learning from AF447: Human-machine interaction', *Safety Science,* 112, pp. 48–56. doi: https://doi.org/10.1016/j.ssci.2018.10.009.

Dehais, F., Behrend, J., Peysakhovich, V., Causse, M. and Wickens, C. D. (2017) 'Pilot Flying and Pilot Monitoring's Aircraft State Awareness During Go-Around Execution in Aviation: A Behavioral and Eye Tracking Study', *The International Journal of Aerospace Psychology,* 27(1-2), pp. 15–28. doi: https://doi.org/10.1080/10508414.2017.1366269.

Delliaux, S., Delaforge, A., Deharo, J.-C. and Chaumet, G. (2019) 'Mental Workload Alters Heart Rate Variability, Lowering Non-linear Dynamics', *Frontiers in Physiology,* 10. doi: https://doi.org/10.3389/fphys.2019.00565.

DiDomenico, A. and Nussbaum, M. A. (2008) 'Interactive effects of physical and mental workload on subjective workload assessment', *International Journal of Industrial Ergonomics,* 38(11), pp. 977–983. doi: https://doi.org/10.1016/j.ergon.2008.01.012.

DiDomenico, A. and Nussbaum, M. A. (2011) 'Effects of different physical workload parameters on mental workload and performance', *International Journal of Industrial Ergonomics,* 41(3), pp. 255–260. doi: https://doi.org/10.1016/j.ergon.2011.01.008.

Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A. R. and Miyake, S. (2016) 'Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study', *Applied ergonomics,* 52, pp. 95–103. doi: https://doi.org/10.1016/j.apergo.2015.07.009.

Fernández-Morales, C., Espejo-Antúnez, L., Clemente-Suárez, V., Tabla-Hinojosa, F. and Albornoz-Cabello, M. (2024) 'Analysis of heart rate variability during emergency flight simulator missions in fighter pilots', *BMJ Mil Health,* 170(4), pp. 296–302. doi: https://doi.org/10.1136/military-2022-002242.

Gaillard, A. W. K. (1993) 'Comparing the concepts of mental load and stress', *Ergonomics,* 36(9), pp. 991–1005. doi: https://doi.org/10.1080/00140139308967972.

Gonzalez, K., Sasangohar, F., Mehta, R. K., Lawley, M. and Erraguntla, M. 'Measuring fatigue through heart rate variability and activity recognition: A scoping literature review of machine learning techniques'. *Proceedings of the human factors and ergonomics society annual meeting*: SAGE Publications Sage CA: Los Angeles, CA, 1748–1752.

Gordan, R., Gwathmey, J. K. and Xie, L. H. (2015) 'Autonomic and endocrine control of cardiovascular function', *World J Cardiol,* 7(4), pp. 204–14. doi: https://doi.org/10.4330/wjc.v7.i4.204.

Green, K. T., Dennis, P. A., Neal, L. C., Hobkirk, A. L., Hicks, T. A., Watkins, L. L., Hayano, J., Sherwood, A., Calhoun, P. S. and Beckham, J. C. (2016) 'Exploring the relationship between posttraumatic stress disorder symptoms and momentary heart rate variability', *Journal of Psychosomatic Research,* 82, pp. 31–34. doi: https://doi.org/10.1016/j.jpsychores.2016.01.003.

Hart, S. G. and Hauser, J. R. (1987) 'Inflight application of three pilot workload measurement techniques', *Aviation, space, and environmental medicine*.

Hirt, C., Eckard, M. and Kunz, A. (2020) 'Stress generation and non-intrusive measurement in virtual environments using eye tracking', *Journal of Ambient Intelligence and Humanized Computing,* 11(12), pp. 5977–5989. doi: https://doi.org/10.1007/s12652-020-01845-y.

Hoover, A., Singh, A., Fishel-Brown, S. and Muth, E. (2012) 'Real-time detection of workload changes using heart rate variability', *Biomedical Signal Processing and Control,* 7(4), pp. 333–341. doi: https://doi.org/10.1016/j.bspc.2011.07.004.

Hörmann, H.-J. 'Training of aircrew decision making'. *AGARD*, 15.

Immanuel, S., Teferra, M. N., Baumert, M. and Bidargaddi, N. (2023) 'Heart rate variability for evaluating psychological stress changes in healthy adults: A scoping review', *Neuropsychobiology,* 82(4), pp. 187–202. doi: https://doi.org/10.1159/000530376.

Järvelin-Pasanen, S., Sinikallio, S. and Tarvainen, M. P. (2018) 'Heart rate variability and occupational stress—systematic review', *Industrial health,* 56(6), pp. 500–511. doi: https://doi.org/10.2486/indhealth.2017–0190.

Kirwan, B. (2024) 'The Impact of Artificial Intelligence on Future Aviation Safety Culture', *Future Transportation,* 4(2), pp. 349–379. doi: https://doi.org/10.3390/futuretransp4020018.

Lee, Y.-H. and Liu, B.-S. (2003) 'Inflight workload assessment: Comparison of subjective and physiological measurements', *Aviation, space, and environmental medicine,* 74(10), pp. 1078–1084.

Lennartsson, A.-K., Jonsdottir, I. and Sjörs, A. (2016) 'Low heart rate variability in patients with clinical burnout', *International Journal of Psychophysiology,* 110, pp. 171–178. doi: https://doi.org/10.1016/j.ijpsycho.2016.08.005.

Li, W.-C., Harris, D., Hsu, Y.-L. and Wang, T. (2011) 'Understanding pilots' cognitive processes for making in-flight decisions under stress'.

Martins, A. P. G. (2016) 'A review of important cognitive concepts in aviation', *Aviation,* 20(2), pp. 65–84. doi: https://doi.org/10.3846/16487788.2016.1196559.

McCraty, R. and Shaffer, F. (2015) 'Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health risk', *Glob Adv Health Med,* 4(1), pp. 46–61. doi: https://doi.org/10.7453/gahmj.2014.073.

Meshkati, N. (1988) 'Heart rate variability and mental workload assessment', *Advances in psychology*: Elsevier, pp. 101–115.

Miller, S. (2001) 'Workload measures', *National Advanced Driving Simulator. Iowa City, United States*.

Miyake, S. (2001) 'Multivariate workload evaluation combining physiological and subjective measures', *International Journal of Psychophysiology,* 40(3), pp. 233–238. doi: https://doi.org/10.1016/S0167-8760(00)00191-4.

Myrtek, M., Fichtler, A., Strittmatter, M. and Brügner, G. (1999) 'Stress and strain of blue and white collar workers during work and leisure time: Results of psychophysiological and behavioral monitoring', *Applied ergonomics,* 30(4), pp. 341–351. doi: https://doi.org/10.1016/s0003-6870(98)00031-3.

Nagasawa, T. and Hagiwara, H. 'Workload induces changes in hemodynamics, respiratory rate and heart rate variability'. *2016 IEEE 16th international conference on bioinformatics and bioengineering (BIBE)*: IEEE, 176–181.

O'Donnell, R. D. and Eggemeier, F. T. (1986) 'Workload assessment methodology', *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance*. Oxford, England: John Wiley & Sons, pp. 1–49.

Parasuraman, R., Mouloua, M. and Hilburn, B. (1999) 'Adaptive aiding and adaptive task allocation enhance human-machine interaction', *Automation technology and human performance: Current research and trends*, pp. 119–123. doi: https://doi.org/10.1177/001872088803000405.

Pham, T., Lau, Z. J., Chen, S. H. A. and Makowski, D. (2021) 'Heart Rate Variability in Psychology: A Review of HRV Indices and an Analysis Tutorial', *Sensors,* 21(12), pp. 3998. doi: https://doi.org/10.3390/s21123998.

R Core Team (2024) *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria*. Available at: https://www.r-project.org/.

Sammito, S. and Böckelmann, I. 'Factors influencing heart rate variability'. *International Cardiovascular Forum Journal*.

Sammito, S., Thielmann, B., Klussmann, A., Deußen, A., Braumann, K.-M. and Böckelmann, I. (2024) 'Guideline for the application of heart rate and heart rate variability in occupational medicine and occupational health science', *Journal of Occupational Medicine and Toxicology,* 19(1), p. 15. doi: https://doi.org/10.1186/s12995-024-00414-9.

Schnell, I., Potchter, O., Epstein, Y., Yaakov, Y., Hermesh, H., Brenner, S. and Tirosh, E. (2013) 'The effects of exposure to environmental factors on Heart Rate Variability: An ecological perspective', *Environmental Pollution,* 183, pp. 7–13. doi: https://doi.org/10.1016/j.envpol.2013.02.005.

Sheridan, T. B. (1995) 'Teleoperation, telerobotics and telepresence: A progress report', *Control Engineering Practice,* 3(2), pp. 205–214. doi: https://doi.org/10.1016/0967-0661(94)00078-U.

Taelman, J., Vandeput, S., Spaepen, A. and Van Huffel, S. 'Influence of Mental Stress on Heart Rate and Heart Rate Variability'. *4th European Conference of the International Federation for Medical and Biological Engineering*, Berlin, Heidelberg, 2009//: Springer Berlin Heidelberg, 1366–1369.

Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-Aho, P. O. and Karjalainen, P. A. (2014) 'Kubios HRV–heart rate variability analysis software', *Computer methods and programs in biomedicine,* 113(1), pp. 210–220.

Tattersall, A. J. and Foord, P. S. (1996) 'An experimental evaluation of instantaneous self-assessment as a measure of workload', *Ergonomics,* 39(5), pp. 740–8. doi: https://doi.org/10.1080/00140139608964495.

Tattersall, A. J. and Hockey, G. R. J. (1995) 'Level of Operator Control and Changes in Heart Rate Variability during Simulated Flight Maintenance', *Human Factors,* 37(4), pp. 682–698. doi: https://doi.org/10.1518/001872095778995517.

The, A.-F., Reijmerink, I., van der Laan, M. and Cnossen, F. (2020) 'Heart rate variability as a measure of mental stress in surgery: A systematic review', *International Archives of Occupational and Environmental Health,* 93(7), pp. 805–821. doi: https://doi.org/10.1007/s00420-020-01525-6.

Thielmann, B., Pohl, R. and Böckelmann, I. (2021) 'Heart rate variability as a strain indicator for psychological stress for emergency physicians during work and alert intervention: A systematic review', *Journal of Occupational Medicine and Toxicology,* 16(1), p. 24. doi: https://doi.org/10.1186/s12995-021-00313-3.

Würfel, J., Djartov, B., Papenfuß, A. and Wies, M. (2023) 'Intelligent pilot advisory system: The journey from ideation to an early system design of an AI-based decision support system for airline flight decks', *AHFE 2023*. doi: https://doi.org/10.54941/ahfe1003844.

Yeh, Y.-Y. and Wickens, C. D. (1988) 'Dissociation of Performance and Subjective Measures of Workload', *Human Factors,* 30(1), pp. 111–120. doi: https://doi.org/10.1177/001872088803000110.