A ROADMAP FOR ASSESSING HUMAN FACTORS INFLUENCE ON PILOT PERFORMANCE BASED ON SIMULATED FLIGHTS

Guilherme Conceição Rocha*, Jens Alfredson**, Luís Gonzaga Trabasso***, Alex Nogueira de Almeida***, Aline da Conceição Matheus***, Diego Hernandez Arjoni***, Manuel Alejandro Rodriguez Diaz***

*Konatus, **SAAB, ***ITA

Abstract

This paper describes the utilization of large amount of data to analyze the performance of pilots under different flight circumstances assessing the influence of human factors like workload and pilot experience.

The methodological approach described herein consists of establishing main research questions, defining hypothesis, acquiring and analyzing simulated flight data. Methods include subjective and objective measurements of qualitative and quantitative nature.

After declaring the research questions and hypothesis, simulated flights were performed and analyzed. The preliminary results inspired the creation of a roadmap for assessing Human Factors influence on pilot performance based on simulated flights.

1 Introduction

The Flight Safety Foundation points human errors as the main cause of aircraft accidents nowadays [1]. As the aircraft system is designed operated and maintained by humans and, according to [2], it is impossible to eliminate human error, a research goal is to understand the mechanisms behind pilot performance and operational errors to develop strategies for performance improvement and error mitigation that can reduce accident rates.

Chapanis [3] defines Human Factors as a body of knowledge about human abilities, human limitations and other human characteristics that are relevant to the design. He also defines Human Factors Engineering as the application of human factors information to the design of tools, machines, systems, tasks, jobs and environment safer, more comfortable and more effective.

In the context of this work, operational safety, comfort and effectiveness are inferred through pilot performance indicators. Therefore, by analyzing how human factors influence pilot performance, one can improve aircraft design, so that accidents rates are reduced, flight operations become more comfortable and missions more effective.

This work is aligned with the initiatives of creating a robotic basis realistic flight simulator at CCM-ITA: the SIVOR platform.

The development of the SIVOR platform motivates the advance of other research perspectives, such as the study of human factors on pilot performance, what corresponds to the scope of the HumAer laboratory.

A bibliographic research was conducted to better organize the research problem of assessing human factors influence on pilot performance, and then to formulate the main research questions.

Workload is a well-established concept, used for evaluation of human work in a wide area of applications, including pilot mental workload. There are numerous methods of assessing mental workload, including psychophysiological, such as heart rate measures [4]. Piloting an aircraft is at times mentally demanding and at times, tasks substantially increase the pilot's mental workload, especially during unexpected events where decisions need to be made quickly. These kinds of situations, although rare, are important to understand since degraded performance may have fatal consequences. Despite pilots being trained in handling extreme situations and aircraft having systems in place to help them, the most contributing factor to accidents is the behavior of the crew. In some incidents and accidents, poor situational awareness or high mental workload has often been a major factor in the sequence of events leading up to the accident [5]. Workload is interesting to study by itself, but it is also interesting to study together with other concepts. The concepts of mental workload and situation awareness (SA) are linked [6]. Mental workload and SA has also been shown to be systematically related to performance [7], [8], [9], [10].

The initial experiments were performed in a simplified fixed-base simulator at a workstation with three different pilots immersed in different flight conditions and subjected to different workload stimuli, following the approach of MATB-II [11], [12]. Workload in flight simulations has been shown to correlate very well with real flight [13].

2 Research Questions

After the bibliographic research, the authors accomplished the following list of research questions:

1. What are the variables or combination of them that best describe pilot performance?
2. Is it possible to induce a realistic stressful situation on a flight simulator, simulating operational environment stress (like ATC communication, envelope edge operations or upset recovery workloads)? How to measure the simulation representativeness?
3. How to estimate the pilot cognitive difficulty to control the flight based on a combination of operational environment and pilot health condition and accumulated experience?
4. How to create operational environment stress indicators (for instance workload indicators)?
5. Can physiological sensors be used to infer the pilot cognitive difficulty (or stress)? How to estimate the pilot cognitive difficulty based on physiological sensors measurement?
6. What are the effects of the workload and other operational environment parameters on the pilot performance?
7. What are the effects of the pilot health condition and accumulated experience on the pilot performance?
8. How physiological sensors influence the operational environment, mainly the pilot performance?
9. What is the influence of the flight simulator visual immersion on the simulation representativeness?
10. What is the influence of the flight simulator motion on the simulation representativeness?
11. What is the influence of the flight simulator inceptors and other human-machine interface on the simulation representativeness?

3 Experimental Procedures Strategies

Depending on the research question, different experimental procedures strategies are applicable to reach the answer to those questions.

The strategies are detailed as follows:

Q1 – Propose different variables that describe pilot performance and verify which ones are more sensitive to human factors parameters.

Q2 – Compare simulated flights with real flights in terms of pilot performance indicators and physiological measurements. Simulation is as more representative as closer is the pilot body response on both simulated and real flights.

Q3 – Propose pilot cognitive difficulty indicators models based on a combination of operational environment and pilot health condition and accumulated experience. Calibrate these models using pilot performance and physiological sensor measurements on simulated flights.

Q4 – Propose operational environment stress models. Calibrate these models using physiological sensor measurements on simulated flights.

Q5 – After defining a simulation environment like real flights in terms pilot behavior measured by physiological sensors, define a combination of those measurements that models the pilot cognitive difficulty and calibrate this model using pilot performance indicators.

Q6 – Perform several simulated flights variating workload and other operational environment parameters and evaluate the influence of them on the pilot performance.

Q7 – Perform several simulated flights with different pilots and with the same pilot at different health conditions (at normal condition, under stress, under the effect of caffeine, with sleep privation, and so on). Verify how the pilot performance is affected.

Q8 – Propose a model for the physiological sensors measurement depending on several factors. Calibrate this model with simulated flight data.

Q9 – Perform several simulations flights with different visual immersion conditions and compare pilot physiological measurements and performance.

Q10 – Perform several simulations flights with different motion conditions and compare pilot physiological measurements and performance.

Q11 – Perform several simulations flights with different simulator inceptors and other human-machine interface conditions and compare pilot physiological measurements and performance.

In addition, both civil and military scenarios have been identified with relevance for the above research questions. The civil scenario has formed the basis for the flight procedure implemented in the study described below. The military scenario includes team collaboration tasks to be performed in reconnaissance missions, for instance, where pilots have to collaborate during events that are initiated either by the external context, such as activities resulting from enemy behavior, as well as limitations in the own action space, such as those due to system degradation/performance or limited resources.

4 Preliminary Experiments

Preliminary experimental procedures were performed to partially answer questions Q1, Q6, Q7 and Q8.

4.1 The Flight Simulator

The experimental procedure was performed with two workstation computers: one handling the flight dynamics model (FDM) and data collection algorithms and the other, processing only the visual rendering task for the pilot.

Fig. 1. Simulation Aircraft used for the experiment.

The FDM used was the JSBSim, which is contained within the FlightGear Flight Simulator software. The simulated aircraft used was a Boeing 737-300 as shown in Fig. 1. Two instances of this simulator were executed; one in the “Control” Workstation handling the FDM, and the other handling the “Visual” task, rendering the virtual scenario and instruments for the pilot.

Data acquisition and experiment control was performed in the “Control” workstation by the experimenter while the pilot interfaced with the “Visual” workstation, the “Control” sends aircraft spatial position data via UDP (Universal Data Protocol) communication to the “Visual” workstation, which displays the information needed by the pilot for decision making. The pilot’s actions are received through a Joystick whose input is feed to the “Control” workstation that process it in the FDM and the flow cycle repeats at a constant rate of 60 Hz as depicted in Fig. 2.

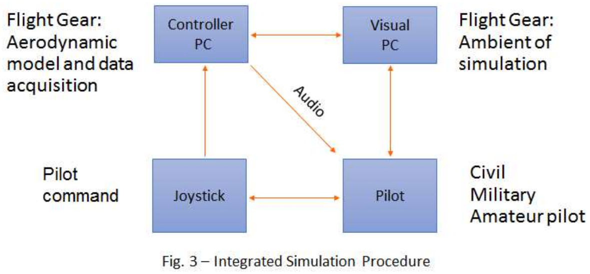

Fig. 2. Data flow chart.

4.2 The Pilots Sample

Three categories of pilots participated in this preliminary experiment, as follows:

- Military aviation pilot (Pilot A).
- Civil commercial aviation pilot (Pilot B).
- Amateur pilot (Pilot C).

An amateur pilot is defined as a person who is not a certified pilot, but has some knowledge of aeronautics, flight simulation experience and is familiar with subjects regarding aircraft operation, whose skill at flight simulation and the basics of aircraft operations are considered sufficient for keeping basic flight tasks in a flight simulator, such as maintaining straight-and-level flight and performing basic maneuvers and procedures but lacking the deep technical knowledge to completely operate a real aircraft on her or his own.

4.3 The Workload Evaluation

Two types of workload were used in the experiment: Primary Workload and Secondary Workload.

Primary workload refers to an inherent added difficulty in maintaining authority over the aircraft; these were designed to require the pilot to exert constant active control over the aircraft or else it will eventually enter an incontrollable situation and would result in a crash.

Secondary workload refers to an occurrence that punctually requires the pilot's attention, and that can be performed without completely interrupting the execution of the primary workload counteractions, requiring enough brief attention to be hindrances to this process.

There are 4 levels of primary workload in this experiment, only one of which occurs during each flight; those are classified by the following types:

Type 1, Normal Flight: no Primary workload occurs.

Type 2, Altered CG: During take-off rotation the aircraft CG will be moved aft to a point where the aircraft longitudinal stability decreases.

Type 3, Engine failure: The right engine will suddenly lose partially its power at a speed of 135 knots during take-off.

Type 4, Aileron trim failure: The aileron trim will suddenly fail and will move to 50% of its travel range at an altitude of 2400ft ASL (Approximately 650ft from the ground at airport);

Secondary workloads are related to warnings (aural and visual) and ATC (Air Traffic Control) instructions. Two types are used in the experiment:

Punctual Alarm: A visual and aural alarm that occurs suddenly in clear view of the pilot, and that need to be deactivated as soon as possible once it occurs by pressing a button, it is intended to reproduce the "Master Warning" system present in most aircraft.

Radio Command: An ATC aural instruction issued to the pilot that need to be complied with as soon as possible. In this case, it requires the pilot to switch the current Communications radio frequency to the one dictated by the instruction.

These two types of secondary workload were combined, creating a total of three possible levels of secondary workload levels.

Besides of them, a workload gap scenario with no secondary workload were also included on the experiment. Thus, four different secondary workload scenarios were evaluated in all flights:

- Workload Gap (only primary workload).
- Primary Workload Disturbance + Punctual Alarm.

- Primary Workload Disturbance + Radio Command.
- Primary Workload Disturbance + Punctual Alarm + Radio Command.

4.4 Flight Procedure

All pilots were instructed in the beginning of the experiment to execute the same flight plan..

The flight profile is displayed in Fig. 3.

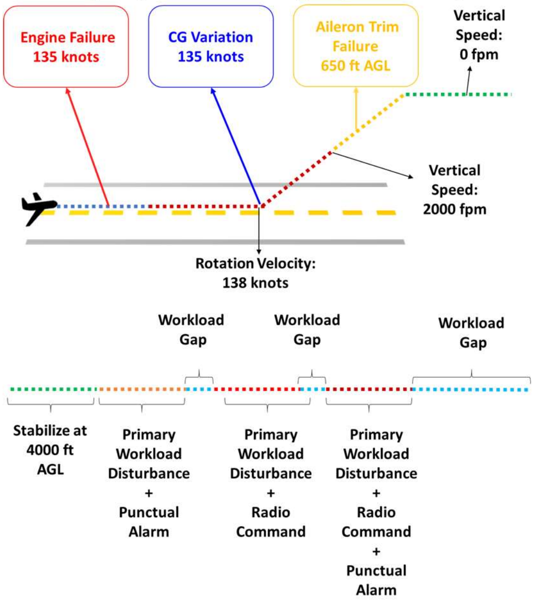

Fig. 3. Flight profile.

The pilot reception follows the flow chart of Fig. 4 and is detailed in [14].

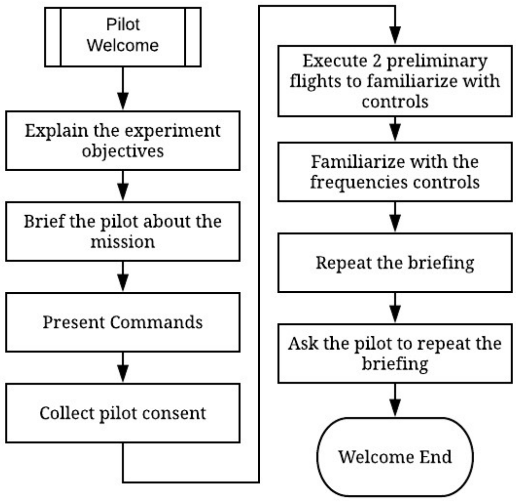

Fig. 4. Pilot reception flow chart.

Each pilot was briefed to:

1. Wear the heart hate meter and position himself at the fixed base simulator;

2. Wait for the command to start the flight procedure;

3. Set the aircraft power to 100%, set flaps to second level, and wait for stabilization of the engines;

4. Release the parking breaks and start the take-off run once engines are stabilized ;

5. Retreat the flaps and landing gear at the altitude of 3000 ft;

6. Execute a rotation of the aircraft as soon as the airspeed reaches 138 knots and keep a vertical velocity of 2000 fpm until 3900 ft is reached;

7. Stabilize the aircraft at 4000 ft with vertical velocity of 0 fpm. Pilot is instructed to keep altitude between 3900 ft and 4100 ft;

8. Keep the heading and the bank angle of the aircraft between -5º and +5º, controlling the aircraft only by the side stick, without changing the power of the aircraft.

9. After stabilizing the aircraft at 4000 ft with 0º of heading and banking angle, the pilot is instructed to change the COMM standby frequency whenever demanded by ATC, and set it to the selected frequency as fast as possible;

10. During simulation, if any alarm turns on, the pilot is instructed to turn it off immediately;

11. At the end of the run, the flight will be paused.

The experimental procedure followed by the experimentation team is illustrated by the flow chart of Fig. 5.

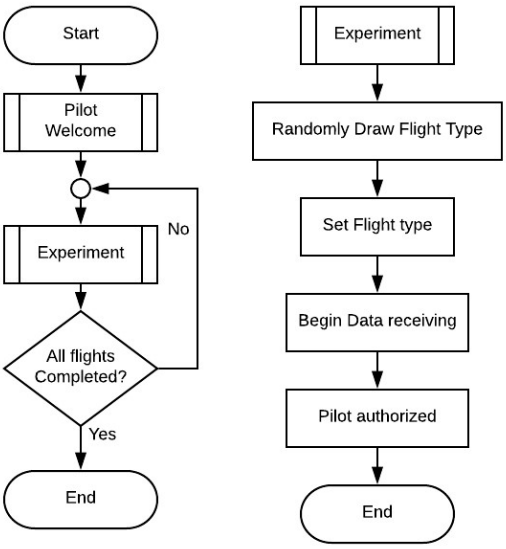

Fig. 5. Experimental procedure.

The experiments are conducted with the pilots wearing a heart rate meter. The heart rate measurement was recorded during all simulated conditions.

4.5 Final Considerations

Five different pilot performance indicators were evaluated:

- Altitude during cruise;
- Rate of climb;
- Heading;
- Yaw rate;
- Mean command intensity.

The following human factors influence on pilot performance were studied:

- Primary workload (or failure);
- Secondary workload;
- The pilot itself (encompassing experience and health condition).

The pilot mental workload was assessed by a heart rate meter.

4.6 Statistical Modeling

As detailed in [14], the statistical model is described by equation (1):

$$V_{ij} = \mu + W_i + P_j + (WP)_{ij} + e_{ij} \qquad (1)$$

where:

V_{ij}: output value;

μ: general output mean;

W_i: variance of the Workload Factor;

P_j: variance of the Pilot Factor;

WP_{ij}: variance of the interaction between the Workload Factor and the Pilot Factor;

e_{ij}: random error.

The considered output values are the five different pilot performance indicators described in section 4.5:

- Altitude during cruise;
- Rate of climb;
- Heading;
- Yaw rate;
- Mean command intensity.

5 Results

Table 1 summarizes the experiment results based on P-values. The pilot performance measurements were marked in light green whenever the criteria for using ANOVA, based on the Shapiro-Wilk test for normality and the Bartlett´s test of homogeneity of variances of the secondary workload.

Table 1. Results summary.

Primary Workload	Pilot Performance	Workload Influence	Pilot Influence	Interaction Influence
Normal	Altitude	0,324	0,0682	0,1939
	Rate of Climb	0,123	1,81E-07	0,329
	Heading	0,210568	0,000346	0,140895
	Yaw Rate	0,4169	0,0147	0,3506

	Command Intensity Indicator	0,705491	0,000211	0,827167
CG	Altitude	0,3538	0,0828	0,3898
	Rate of Climb	0,0466	0,0332	0,4979
	Heading	0,050274	0,000124	0,338794
	Yaw Rate	0,0771	0,0333	0,1559
	Command Intensity Indicator	0,677	2,63E-08	0,819
Motor	Altitude	0,0988	0,1247	0,1163
	Rate of Climb	0,0127	6,80E-07	0,3439
	Heading	0,277	3,14E-08	0,426
	Yaw Rate	0,13685	0,9834	0,00664
	Command Intensity Indicator	0,6717	0,0996	0,7686
Aileron Trim	Altitude	0,174	0,125	0,327
	Rate of Climb	0,000629	1,16E-05	0,017196
	Heading	0,107	3,84E-09	0,193
	Yaw Rate	0,4102	0,0279	0,7637
	Command Intensity Indicator	0,4349	0,0477	0,7537

Based on figures above, using a confidence factor of 95%, one can conclude that the outstanding variable is the rate of climb, in that this is the only indicator that is sensitive to the secondary workload variation.

It can also be noticed that both the secondary workload variation as well as the pilot influence the precision of the rate of climb.

Finally, Table 2, Table 3, Table 4., and Table 5 summarize the maximum variation noticed on the heart rate meter on several primary and secondary workload conditions.

Table 2. Type 1, Normal Flight.

	W1	W2	W3	W4
PILOT A	8,657	1,657	2,657	4,657
PILOT B	9,053	1,053	7,053	9,053
PILOT C	0,535	3,464	7,535	1,735

Table 3. Type 2, Altered CG.

	W1	W2	W3	W4
PILOT A	5,657	5,657	2,657	2,657
PILOT B	3,053	3,053	2,253	5,053
PILOT C	3,701	7,535	10,535	11,535

Table 4. Type 3, Engine Failure.

	W1	W2	W3	W4
PILOT A	8,657	5,257	6,657	1,657
PILOT B	3,053	1,053	2,053	2,053
PILOT C	3,535	1,393	7,535	0,264

Table 5. Type 4, Aileron Trim Failure.

	W1	W2	W3	W4
PILOT A	9,657	4,657	9,323	5,657
PILOT B	9,053	1,053	3,053	1,053
PILOT C	1,535	1,464	5,535	4,535

Based on the previous results, it is possible to notice that only the pilot C (amateur pilot) was sensitive to the variation of secondary workload at altered aircraft CG flight condition.

The plot in Fig 6. depicts the correlation between secondary workload and hart rate variation for the amateur pilot on the altered CG condition. To build this graphic a secondary workload index was developed varying from 1 to 4, incrementally.

It is important to highlight that this correlation was not observed in other scenarios and a deeper study with more test data need to be performed to elucidate the reasons for that.

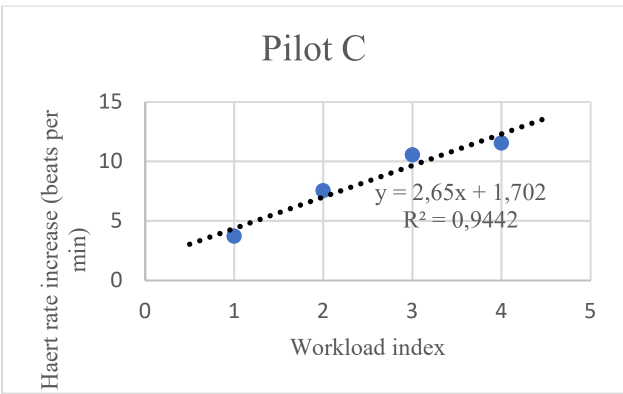

Fig. 6. Heart rate at various levels of workload for Pilot C, the amateur pilot.

6 Conclusions

The research questions defined in this paper establish a roadmap for conducting experiments with simulated and real flights to assess the influence of human factors on pilot performance.

The preliminary experiments results demonstrated that only a few performance indicators are sensitive to secondary workload and more tests and studies need to be conducted to better understand this phenomenon.

Pilot background influence on the performance was confirmed, and the way the physiological measurement of the heart rate varies with the secondary workload change dramatically depending on the pilot and on the primary workload condition.

Although the preliminary experiments provide some insights on how human factors influence the pilot performance during the studied conditions, additional tests, research and analysis are still necessary to advance more on the answer of the research questions proposed.

7 Acknowledgements

This study is part of an ongoing cooperation regarding integrated vehicle health management and human factor analysis. This research has been supported by the Swedish Governmental Agency for Innovation Systems (Vinnova) and Financiadora de Estudos e Projetos (Funding Authority for Studies and Projects, FINEP).

References

1. Flight Safety Foundation Editorial Staff. Flight Crew Procedures Streamlined For Smoke/Fire/Fumes. Flight Safety Foundation. Flight Safety Digest, June (2005)
2. Dekker, S. The field guide to understanding human error. Burlington, VT: Ashgate Publishing Company, (2006)
3. Chapanis, A. To Communicate the Human Factors Message, You Have to Know What the Message Is and How to Communicate It. Human Factors Society Bulletin, November (1991)
4. Castor, M., Hanson, E., Svensson, E., Nählinder, S., LeBlaye, P., MacLeod, I., Wright, N., Alfredson, J., Agren, L., Berggren, P.: GARTEUR Handbook of Mental Workload Measurement. Grouo for Aeronatical Research and Technology in Europe, Flight Mechanics Action Group FM AG 13 (2003)
5. Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F.: Measuring Neurophysiological Signals in Aircraft Pilots and Car Drivers for the Assessment of Mental Workload, Fatigue and Drowsiness. Neuroscience & Biobehavioral Rev. 10(44), 58-75 (2012)
6. Wickens, C. D. Situation awareness and workload in aviation. Current Directions in Psychological Science, 11(4), 128-133. (2002)
7. Castor, M. The use of structural equation modeling to describe the effect of operator functional state on air-to-air engagement outcomes. Doctoral dissertation (No. 1251). Linköping, Sweden: Linköping University. (2009)
8. Svensson, E., Angelborg-Thanderz, M., & Sjöberg, L. Mission challenge, mental workload and performance in military aviation. Aviation, Space, and Environmental Medicine, 64(11), 985-991. (1993)
9. Svensson, E., Angelborg-Thanderz, M., Sjöberg, L., & Olsson, S. Information complexity: Mental workload and performance in combat aircraft. Ergonomics, 40(3), 362-380. (1997)
10. Svensson, E., & Wilson, G. F. Psychological and psychophysiological models of pilot performance for systems development and mission evaluation. The International Journal of Aviation Psychology, 12(1), 95-110. (2002)
11. Hart, B., Stavenland, L.E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Human Mental Workload, 139-183. Elsevier (1988)
12. Comstock Jr, J. R., & Arnegard, R. J. MATB - Multi-AttributeTaskBatteryfor human operator workload an d strategic behavior research Technical Memorandum 104174. Hampton, Virginia: NASA, Langley Research Center. (1992)
13. Nählinder, S. Flight Simulator Training: Assessing the Potential. Doctoral dissertation (No. 1250). Linköping, Sweden: Linköping University (2009)
14. Arjoni DH et al. Artigo ICAS. Experimental analysis of flight performance under workload variations. The 31st Congress of the International Council of the Aeronautical Sciences. Belo Horizonte, Brazil: ICAS (2018).

Copyright Statement