Experimental Evaluation of the Human Performance on a RoboticFlight Simulator based on FOQA Parameters

D. H. Arjoni, G. C. Rocha, A. H. Moreira, R. M. Nicola, W. R. Oliveira, A. V. S. Silva, G. S. Natal, L. Silveira, E. Thomas, E. Villani and Luís Gonzaga Trabasso

Book Chapter

Part of: Proceedings of the Aerospace Technology Congress. Kaj Lundahl, Roland Karlsson, Björn Jonsson and Knut Övrebö (eds) , 2016, pp. 1-11.

TECHNICAL SESSION: ***Aircraft and spacecraft system analysis***

TOPICS: Modelling & simulation, concept development

Experimental Evaluation of the Human Performance on a Robotic Flight Simulator based on FOQA Parameters

D. H. Arjoni [1], G. Rocha[2], R. Pereira[2], A. H. Moreira[1], R. M. Nicola[1] ,W. R. de Oliveira [1], A. V. S. Silva[1], G. S. Natal[1], L. Silveira[4], E. Thomaz[3], E. Villani[1], L. G. Trabasso[1]

[1]: ITA Instituto Tecnológico de Aeronáutica, São José dos Campos - SP-Brazil.
[2] Konatus, São José dos Campos - SP – Brazil.
[3] Embraer, São José dos Campos – SP - Brazil.
[4]ANAC, São José dos Campos - SP Brazil

diegoarjoni@gmail.com, guilherme.rocha@konatus.com.br, raphael.pereira@mail.konatus.com.br, anderson.hmoreira@gmail.com, rnicola@outlook.com, weslyro@ita.br, silvaandrevs@gmail.com, guilesn@gmail.com, , leandro.silveira@anac.gov.br, edmar.thomaz@embraer.com.br, evillani@ita.br, gonzaga@ita.br.

The SIVOR project, currently being developed by ITA and Embraer, consists of designing and implementing a high fidelity flight simulator based on the use of COTS industrial robots. The aim of the project is to provide a cost-efficient and flexible platform that can be used along the design phases of the aircraft. One of the advantages of an industrial robot over the traditional Stewart platform is the availability of a large workspace, which provides more flexibility for defining the washout filter. This filter converts the aircraft dynamics into robot movements, which has a limited workspace. The main purpose of the flight simulator is to provide a motion feeling similar to the one imposed by the aircraft movements in a real flight. The representativeness of the motion cue is usually evaluated in a qualitative way by the pilots that fly the simulator. Quantitative methods to evaluate the entire range of actuation of a simulator are complex, inducing tests in fractions of the flight to increase performance. In this work, we discuss the use of FOQA (Flight Operational Quality Assurance) as an additional quantitative tool for the evaluation of the motion cue in the SIVOR flight simulator. FOQA is a voluntary safety program from FAA, detailed in AC-120-82. It proposes a set of parameters that can be used by airliners to analyse flight safety and increase operational efficiency. The verification of FOQA parameters checks whether or not the pilot complies with the standard operational procedures defined by the airliners and aircraft manufacturers. The purpose of this work is to analyse whether or not, and to what extent, the FOQA parameters can be used to evaluate the quality of the motion cue of flight simulators. For this purpose, we define an experimental procedure that compares flights performed by pilots under different motion modes. It then calculates a set of behavioural parameters that has been proposed in order to quantify how the motion affects the inputs of the pilot. The results are submitted to ANOVA statistical analysis that verifies the relevance of the motion factor. Finally, we discuss the capability of a FOQA based experiment to estimate the contribution of the motion to the realism of the flight simulation.

INTRODUCTION

One of the main purposes of dynamic flight simulators is to provide a motion feeling similar to the one imposed by the aircraft movements in a real flight (Giordano et al. 2010). One way to induce these sensations is through a washout filter, a control system capable of converting the aircraft workspace into robot movements to simulate the accelerations of a real flight (Grant & Reid, 1997).

The SIVOR project, currently being developed by ITA and Embraer, consists of designing and implementing a high fidelity flight simulator based on the use of an industrial anthropomorphic robot. For the preliminary phases of the project, the prototype in fig 1 is built with the objectives of develop and evaluate control systems such as the washout filter.

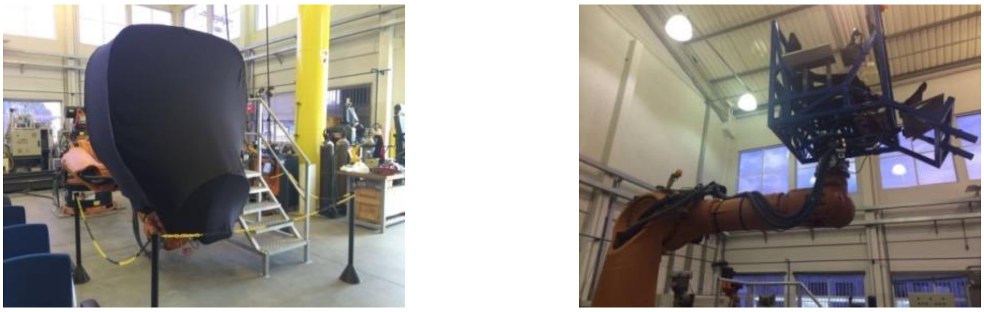

Figure 1 - SIVOR research simulator prototype

The certification and evaluation of flight simulators are usually executed by competent governmental entities such as FAA, ANAC, among others. A great amount of the dynamic evaluation is performed by pilots, whose task is to subjectively assess the quality of the motion cues (14 C.F.R. § 60 2008).

An increasing number of researchers, looking for alternatives to this kind of analysis, use logical methods of experimentation such as expert systems for tuning washout filters (Grant & Reid, 1997). Other research fields are based on the use of the mathematical models of the physiological human perception systems, to assure the quality of the induced sensations (Telban et al. 2005). The use of indicators capable of evaluating the behaviour of the pilot in a flight are suggested by some studies once they are considered to take into account the learning transferring rate achieved in a simulator (Telban et al. 2005), (Grundy et al., 2016).

This work discusses an adaptation of the FOQA Program (Flight Operational Quality Assurance) as an additional quantitative tool for the evaluation of the motion cue control system in the SIVOR flight simulator. FOQA is a voluntary safety program from FAA, detailed in AC-120-82. It proposes a set of parameters which can be used by airliners to analyse flight safety and increase operational efficiency. The verification of FOQA parameters checks whether or not the pilot complies with the safety flight levels defined by the airliners and aircraft manufacturers (US DOT 2004).

METHODS

This section describes the development of the experimental procedure to be executed by pilots, as well as the statistical model and analysis to be applied to the produced data.

Procedure Background

In order to verify the capacity of the FOQA program and the proposed set of behavioural parameters to evaluate the motion of a simulator, an experiment is designed with the requirements of being fast and concise. The test must be held in just a fraction of the flight, once the complete procedure would demand a great amount of time.

The events outlined in the FOQA regulation are able to evaluate a standard flight, sectioned in three distinct phases: take-off, cruise and landing. Each event can be assigned to one or more of the three sections; therefore, each section has its own evaluation group as presented in fig 2.

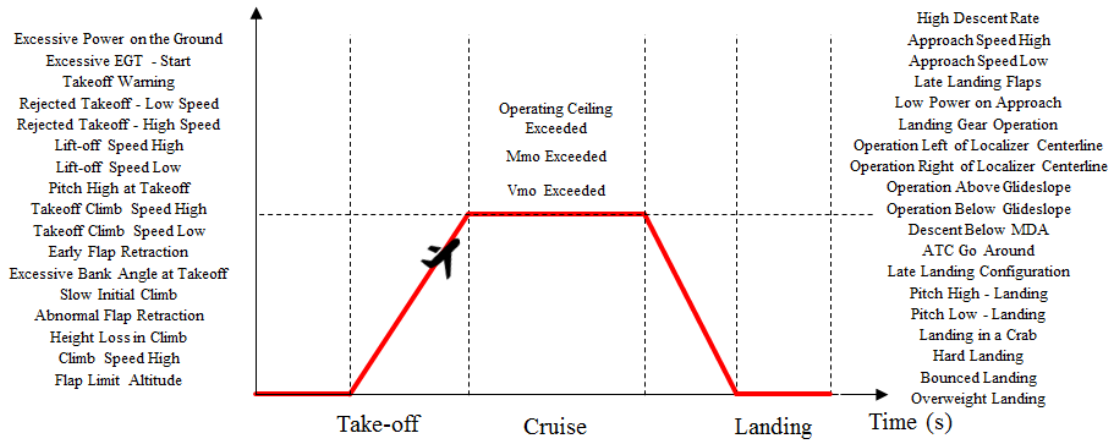

Figure 2 – Flight phases and FOQA events

Considering a standard flight executed without the autopilot, the cruise phase, besides being the longest, does not have a meaningful interaction between pilot and aircraft, and can be eliminated from the analysis along with its FOQA events. The landing phase can be considered the most interactive portion of the flight as well as the most difficult to be executed, but, the addition of this manoeuvre could result in excessive test time due to potential mistakes of the pilots. The suggested experiment is built based on the take-off phase of the flight since it is a straight forward section that can be quickly tested and repeated.

Events related with the aircraft mechanics, ATC interaction or force feedback actuation were eliminated from the analysis, once the measurement of these parameters were out of the scope of this work.

Experiment Metrics

In order to eliminate the binary characteristic from an analysis based exclusively on the FOQA events occurrence, a set of parameters is proposed to quantify the deviation between real human performance and ideal standardized procedures. The behavioural parameters are inspired in the NASA TLX (Task Load Index) tool, a subjective assessment of workload in human-machine based systems (Hart, 1988). The NASA TLX consists of a set of evaluations concerning mental demand, physical demand, temporal demand, performance, effort, and frustration level. Based on the TLX, three objective behavioural parameters are created to measure the pilot performance in the FOQA events:

- Workload: based on the physical demand, mental demand and effort. Evaluated through the absolute, de-trended integral of the pilot input in the joystick axes;
- Precision: based on the performance of the pilot. Evaluated through the mean and standard deviation associated with the attitude variable (roll, pitch or yaw) related to the manoeuvre;
- Response time: based on the temporal demand. Evaluated through the differential time between a stimuli and the actuation command of the pilot, or the summation of a group of actuations originated from a stimuli.

In order to create a flight profile, a new set of events were created based on FOQA and human performance indicators as shown in Table 1.

Table 1 – Human Performance Indicators

N	Event	Description	Workload	Precision	Response Time
1	Reject Takeoff Speed	Verifies the response time and the workload to an engine failure at 100knots	X		X
2	Lift Off Speed	Verifies the reaction time to a 120 knots rotation velocity			X
3	Takeoff Climb Speed	Verifies the workload and the precision in a 140±5 velocity maintenance at the initial climb, until the stabilization at 3000±100 ft altitude from the airport	X	X	
4	Clean up Attitude	Verifies the sum of the response times in the accomplishment of the simultaneous tasks : Flap retraction, Landing gear retraction, power reduction to 60% and altitude stabilization at 3000±100 ft.			X
5	Bank Angle Analysis	Precision and workload are measured in the maintenance of 0° of roll.	X	X	
6	Climb Analysis	Precision and workload are measured to the maintenance of 15° pitch angle at 80% of throttle.	X	X	

The last step is to assign a flight procedure that includes the group of new FOQA events as shown in fig 3.

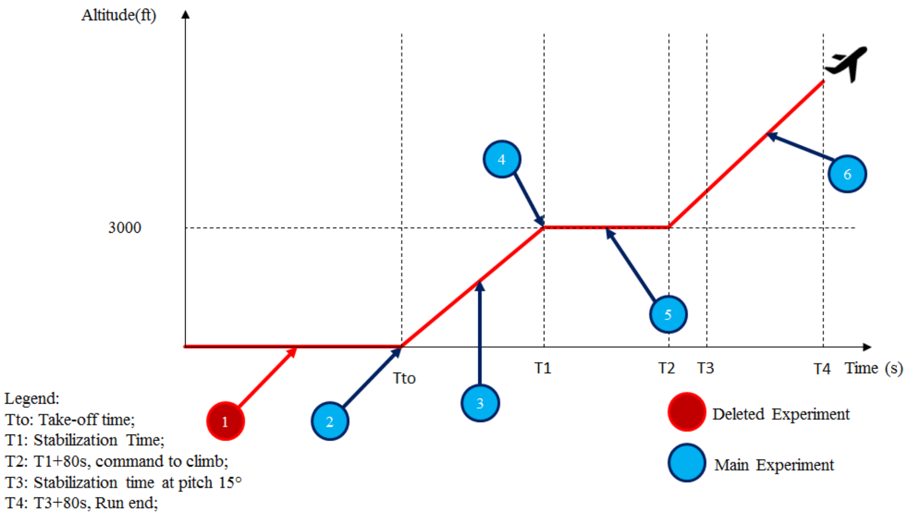

Figure 3 - Flight Procedure

The point 1 is excluded from the main procedure once it does not provide a continuous flight aspect and its execution would lead to the addition of 6 new runs, increasing the setup time and consequently the experiment time.

Statistical Procedure and Analysis

The flight analysis is based on a data frame composed of parameters such as: altitude, attitude, speeds, engine thrust and inputs and outputs of the commands of the pilot. The whole data is subjected to off-line software which analyse the behavioural parameters associated with the events, as described above.

The experiment has been designed to be applied to 7 pilots with flight experience that are subjected to two simulation modes (dynamic simulation and static simulation). In the statistical analysis, the pilot is considered a blocked factor in order to isolate the variance associated with the pilot from the variance related to the simulation mode. The pilot must fly 3 rounds in each mode.

Each variable is subjected to an ANOVA test with significance of 10%. The generic statistical model used for all the outputs is described in Equation (1):

$$V_{ij} = \mu + M_i + \beta_j + e_{ij} \qquad (1)$$

where:

V_{ij}: Output value: mean, standard deviation, workload or response time;

μ: General output mean;

M_i: Simulation mode variance;

β_j: Pilot block variance;

e_{ij}: Random error variance.

The complete output set is composed of 11 variables, associated with the behavioural parameters and the flight events, as described in Table 2.

Table 2 - Variables Characteristics

Event	*Variable Name*	*Behavioural Parameter*
Lift Off Speed	TR120	Response Time
Takeoff Climb Speed	MS	Precision: Mean
	VS	Precision: Standard Deviation
	CS	Workload
Clean up Attitude	TT	Response Time
Bank Angle Analysis	M_sA	Precision: Mean
	V_sA	Precision: Standard Deviation
	C_sA	Workload
Climb Analysis	M_sE	Precision: Mean
	V_sE	Precision: Standard Deviation
	C_sE	Workload

Table 3 must be completed for each one of the eleven variables.

Table 3 - Experimetal Table

	Pilot			
Mode	1	2	...	N
Static	*	*	*	*
	*	*	*	*
	*	*	*	*
Dynamic	*	*	*	*
	*	*	*	*
	*	*	*	*

The simplified flux in the fig 4 specifies the execution of the experiment. All the 6 runs of each pilot must be completed in the same day, avoiding physical and psychological variations. In order to set and adapt the pilot to the simulator, 2 preparatory runs are executed. The run order was randomized to compensate the auto training factor.

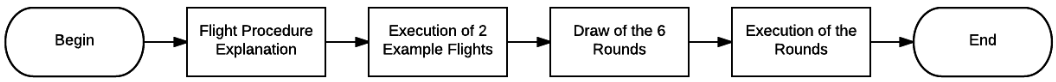

Figure 4 - Simplified experimental procedure

RESULTS AND DISCUSSION

This section displays the results obtained from the statistical analysis and discusses some significant issues concerning both the experiment and the influence of the simulator motion system.

Statistical Analysis

The analysis of the collected data is based on Analysis of Variances. In order to assure its reliability, the sample must have a normal characteristic. For this purpose, two graphical tests are performed: a quantil-quantil graph and a residual dispersion plot. An additional Shapiro-Wilk test mathematically confirms the normality of the sample.

The collected data is naturally susceptible to outliers which prevention is complex; thereby the wrong points are excluded from the sample, without a significant harm.

The boxplot data of all outputs are shown in the fig 5.

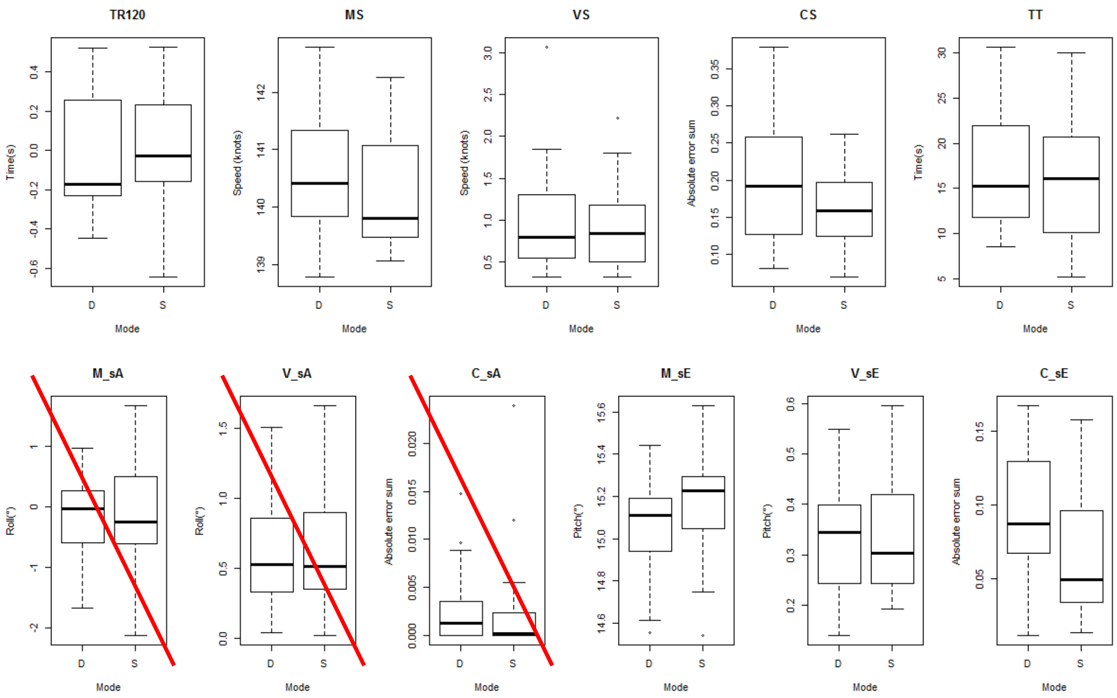

Figure 5 – Boxplots

Table 4 shows the ANOVA test results through its P-Values, confirming or rejecting the analysis of variance hypothesis.

Table 4 - Experiment Results

Variable	Residuals Df	P-Value	Significance
TR120	32	0.535	Not Significant
MS	33	0.171	Not Significant
VS	33	0.547	Not Significant
CS	33	0.005	Significant
TT	30	0.311	Not Significant
M_sA	-	-	Invalid
V_sA	-	-	Invalid
C_sA	-	-	Invalid
M_sE	33	0.0201	Significant
V_sE	33	0.8051	Not Significant
C_sE	33	0.0088	Significant

The Bank Angle Analysis consisted on maintaining a roll angle of 0° and a great number of pilots could perform this task without a significant interaction with the inceptors, causing an tendency on the workload data and consequently in the precision. To avoid misinterpretations from the analysis, the M_sA, V_sA and C_sA were excluded from the possible outputs.

CONCLUSIONS

Starting from the hypothesis that the motion of a flight simulator alters the human performance, the results show that a test based on FOQA events and behavioural parameters is not able to confirm that the motion of a flight simulator has a significant effect in several measurements of the human performance, once, just three of the created variables (CS, M_sE and C_sE) were sensitive to the motion mode used.

A significant variation of the workload was observed during the Takeoff Climb Speed and in the Climb Analysis (CS, C_sE), where the motion difficult the execution of the task, represented by a workload increase. Since the workload in absolute terms is small, the increase observed in the variable is also small.

The sensibility of mean variable referent to the precision in the Climb Analysis (M_sE) to the variation of the simulation mode can be explained by an increase of the inclination perception by the motion, causing the pilot to react more rapidly to the pitch stimuli which leads to an increase of both workload and precision.

Although the sensations created by the motion in a flight simulator are clearly perceived by the pilots, as noticed on a qualitative evaluation of the increased realism of the motion, the current experiment was not able to reflect significant variations on the human performance measurements due to this increased realism. This observation can be explained by some factors, such as:

- The manoeuvres tested were not able to stress enough the motion dynamic once it was based in a standard flight.
- The role of the motion in the tested manoeuvres was not important enough to completely modify the pilot's response.

Based on these information, a possible future work is to develop a testing method capable of comparing the physiological sensations felt in a real flight with the ones induced in the simulation, both with and without motion, although being a more complex experiment, it can be more conclusive once the data has to be collected before the pilot's physical actuation.

ACKNOWLEDGEMENT

The authors acknowledge the financial support of FAPESP (Processes 2012/51085-3 and 2015/03832-2), CNPq (Processes 132801/2015-9 and 306408/2014-7) and CAPES.

REFERENCES

Journals

Hart, S. G., Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index) Results of Empirical and Theoretical Research. Human Mental Workload, p. 139–183.

Grant, P.R., Reid, L. D. (1997). Motion Washout Filter Tuning: Rules and Requirements. Journal of Aircraft: 145-151.

Grant, P.R., Reid, L. D. (1997). PROTEST: An Expert System for Tuning Simulator Washout Filters. Journal of Aircraft, p. 152–159.

Grundy, J. G. et al. (2016). The Effectiveness of Simulator Motion in the Transfer of Performance on a Tracking Task is Influenced by Vision and Motion Disturbance Cues. Human Factors, vol. 58, p. 546–559.

Conferences

Robuffo, G. P., Masone, C., Tesch, J., Pollini, L., Bulthoff, H., H. (2010). A Novel Framework for Closed-Loop Robotic Motion Simulation - Part I: Inverse Kinematics Design. Proceedings of the 2010 IEEE International Conference on Robotics and Automation (2010): 3876-883.

Books

Montgomery, D. C. (2013) Design and Analysis of Experiments. Hoboken, NJ: John Wiley &Amp; Sons, Inc.

Technical Reports

Flight Simulation Training Device Initial and Continuing Qualification and Use 14 C.F.R. § 60 (2012).

Telban, R. J., Cardullo, F. M. (2005) Motion Cueing Algorithm Development: Human-Centered Linear and Nonlinear Approaches. Rep. Hampton: NASA.

Telban, R. J., Cardullo, F. M., Kelly, L. C. (2005). Motion Cueing Algorithm Development: Piloted Performance Testing of the Cueing Algorithms. Rep. Hampton: NASA.

US DOT (2004): Advisory Circular: Flight Operations Quality Assurance (FOQA), AC No: 120-82.