AHFE International

The Application of Human Factors in Wake Vortex Encounter Flight Simulations for the Reduction of Flight Upset Risk and Startle Response

Frédéric Rooseleer[1], Barry Kirwan[2], Diana Paola Moreno Alarcon[3], Elizabeth Humm[4], and Gianluca Borghini[5]

[1]European Organization for the Safety of Air Navigation (EUROCONTROL), Rue de la Fusée 96 1130 Brussels, Belgium
[2]EUROCONTROL Experimental Centre (EEC), Brétigny/Orge, France
[3]Ecole Nationale de l'Aviation Civile (ENAC), Toulouse, France
[4]DeepBlue, Rome, Italy
[5]Sapienza University of Rome, Rome, Italy

ABSTRACT

The current top safety risk concern for commercial air travel in Europe is known as "Flight Upset". This term, also known as "Loss of Control in Flight", entails the flight crew suddenly finding themselves in an unexpected, complex, and even confusing situation, which might result in an accident, possibly fatal. An undesirable aspect of such events is known as the "startle response", wherein one or both flight crew, suddenly finding themselves in confusing and dangerous conditions, may experience 'startle', which temporarily affects their cognitive functioning. A Horizon 2020 research project called SAFEMODE, which aims to integrate Human Factors techniques into a unified framework for designers in aviation and maritime domains, is exploring the use of state-of-the-art flight simulation facilities to measure pilot performance in severe wake turbulence events, which can induce the startle effect. This is part of a broader use case within SAFEMODE to validate the design of a new Wake Vortex Air Traffic Alert for the Cruise phase of flight. The cockpit flight simulations involve type-rated flight crews in realistic and representative cruise flight conditions, using a full-flight motion-based simulator. During the simulations, pilots are exposed to significant wake vortex encounters. The measures selected for analyzing performance in the cockpit flight simulations include Eye Tracking and Electro-Dermal Activity, flight parameters, expert observations and subjective pilot feedback using various scales for workload, situation awareness etc. This paper outlines the Human Factors approach taken, and the preliminary insights gained from the first set of simulation tests.

Keywords: Flight upset, Cockpit simulator training, Startle effect, Human performance monitoring, Wake vortex

INTRODUCTION

The development of new operational concepts needs to adequately integrate Human Factors (HF). A Horizon 2020 research project called SAFE-MODE aims to support such integration via a HUman Risk-Informed Design (HURID) framework. Rather than take a piecemeal approach, focusing on either Human Factors, learning from incidents, or risk modelling, SAFE-MODE integrates all of these into one framework. SAFEMODE also seeks cross-fertilization of the approach to human factors between two different transport modes, via a consortium of Partners from the maritime and aviation domains. HURID is a human-centered approach, requiring analysis of Human Factors with the involvement of end-users such as operational experts, designers, Human Factors experts, and safety experts, as well as simulations (from low to high fidelity) and risk modelling. One of the SAFEMODE aviation case studies used to validate the HURID framework is developing a new operational concept that aims to facilitate the safe management by Flight Crews of wake turbulence encounter upsets in flight (Rooseleer et al, 2021).

Wake Turbulence Encounter Risk and Startle Effect

Every lift-generating, hence flying, aircraft trails wake vortices. The trailing vortices roll-up into a pair of coherent, counter-rotating vortices that can persist for several minutes after the wake-generating aircraft has flown by, potentially causing a hazard to any following aircraft that may encounter these vortices, resulting in an aircraft upset (induced roll, loss of height or rate of climb), potential loss of control in-flight (LOC-I) and cabin injuries. Despite maintaining the correct traffic separation according to the applicable rules, wake turbulence encounters are being occasionally reported in the En-route / cruise phase of flight in some airspace. Some of these events have resulted in significant upsets (reaching up to 60° bank), in particular for smaller aircraft types such as business jets. Experience has demonstrated that if the pilot reacts at the first roll motion – possibly influenced by the startle effect – when in the core of the vortex, the roll motion could be potentially amplified by this initial piloting action, due to the lack of anticipation. Startle and surprise effects can influence pilot performance in many detrimental ways: from mere distraction to inappropriate actions or hasty decision-making. Well-learned procedures and skills can be discarded and substituted by inappropriate reactions, including freezing or over-reacting at the controls, and can therefore impact performance and endanger safety. Startle response has played a substantial role in a significant number of Loss-of-Control In-flight events as well as in other types of accidents (Field et al, 2018). Flight upset risk from wake vortex encounters in en-route airspace have remained unaddressed except via separation procedures and flight crew standard operating procedures, because until now, en-route Air Traffic Control has had no specific means to detect wake encounter risk.

Wake Alerting Concept

A tactical wake turbulence risk alerting the Flight Crews ahead of the encounter could therefore be beneficial to reduce the startle effect and support

Figure 1: Wake alert displayed to controller.

appropriate management of these conflicts. The envisaged risk-alerting logic relies on a ground-based predictor, connected to the Air Traffic Control system, using information from air traffic surveillance and meteorological conditions in upper airspace, as well as a wake behaviour evolution and an encounter risk model. The wake turbulence encounter alert function (Figure 1) is displayed to the En-route / Area Control Centre (ACC) to both Executive & Planning Air Traffic Controllers (ATCOs). This alert will be calculated and displayed at a time horizon of 2-3 minutes before the predicted wake upset. Since the primary goal for Loss of Control in flight (LOC-I) risk mitigation is to inform the flight crew of the imminent wake encounter risk, the wake alert is sought to be classified as a 'Caution' or 'Advisory' type of alert: for conditions that require immediate flight crew awareness and subsequent flight crew response. Once the ATCO receives the alert, the Executive Controller will inform the flight Crew of the exposed aircraft at the earliest opportunity. The 'Caution' information will be accompanied by traffic information about the generating aircraft. The wake turbulence caution will necessitate a decision-making process by Flight Crews on the subsequent actions to take: either decide for a trajectory adaptation for wake upset prevention, or prepare for a possible (significant) wake upset in accordance with Standard Operating Procedures. This includes securing the cabin (set seatbelt sign ON, Passenger Announcement (if time allows), Cabin Crew instruction such as 'stop service' and 'get seated', and as adjusting the Pilot seat position and 'covering the controls', to be ready for managing the aircraft upset, especially if the degree of roll causes the auto-pilot to switch off.

HUMAN ASSURANCE VALIDATION WITH WAKE UPSET FLIGHT SIMULATION

HURID encompasses the entire design and validation process. Early HURID applications included incident analysis and discussions with pilots who had experienced wake encounters, task analysis, prototyping and HAZOP with controllers and pilots (Rooseleer et al, 2021). This early work prepared the way for real-time simulations to see how flight crew would react to the alert, and whether it assisted in flight management and the reduction of startle response. Such validation of the Wake Alert concept requires high-fidelity cockpit flight simulations, involving licensed flight crews in representative conditions, providing the means to evaluate the human performance of the Flight Crews to prepare for, manage or avoid wake encounter upset and

Figure 2: Flight Simulator (interior).

Figure 3: Flight Simulator (exterior).

resulting safety benefits. The challenge facing the Human Factors specialists on the project was therefore how to validate the concept, and in particular, which Human Factors techniques and measures to employ in the cockpit flight simulations. Before detailing the methods and measures selected, as well as the rationale for those selections, it is useful to outline the flight simulator and the trial conditions, since these represented the environment and practical constraints within which the HF specialists needed to work.

Flight Simulator, Flight Crew and Trial Regime

The cockpit flight simulations involve type-rated flight crews in realistic and representative cruise flight conditions, using a Type VI Boeing 737-800 full flight motion-based simulator which is also used for Upset Prevention and Recovery training programs (Figures 2 and 3). Participating aircrew are licensed commercial line pilots from two separate airlines. All are experienced on the same aircraft type. The simulated en-route / cruise phase of flight wake upset inputs were calibrated against actual wake encounter flight data, with support from a major aircraft manufacturer and Partner in the project. The airline pilots involved in the preparation and test phases considered that the simulated wake upset was very realistic. It is worth noting that one of the experienced pilots did actually experience startle response during the first simulated wake encounter, resulting in this pilot missing communication from the co-pilot, and a sense of being unable to react for fifteen to twenty seconds.

Table 1. Flight profile for all runs.

Parameter	Value
Time of day	Daytime, from dawn to dusk conditions
Phase of flight	Cruise
Aircraft status	Cruise speed / wings level / at various flight levels above FL285 / all system functions normal / no planned change of flight level
Autopilot status	Engaged but disengages when the vortex is encountered

The profile of the flight is set for all runs as follows (Table 1):

During the simulation runs, pilots are exposed to simulated wake vortex encounters, corresponding to a strong wake-induced upset (between 30 and 40 degrees of bank), with or without prior ATC wake caution, and varying the initial direction of roll between left and right to limit the simulation training effect. The test period is made up of a series of 24 runs of 45 minutes each, with three wake vortex encounters (randomized) per run. Each wake encounter scenario will last 15 minutes, in which the crew have approximately 7-10 minutes of straight and level flight before the ATC alert is made (or not in the case of the 'control' encounter) and the wake vortex is encountered. The timings differ slightly between run according to the 'time to onset' instructed in the ATC alert, e.g. 3 minutes, 2 minutes, or 'imminent'.

In order to maximize data collection and to represent a real life situation, each flight crew pair (Captain and Flight Officer) will do four runs together, and will alternate position from 'pilot flying' and 'pilot monitoring' between runs, to counter possible flight crew seniority and experience effect on the results. No special training of aircrew will be conducted in this study, however there will be a short period of trial familiarisation at the start of the run for each new crew in the simulator, so the crew can become familiar with the simulator and its instrumentation.

Experimental Design: Independent Variables & Hypotheses

The principal experimental independent variable is the existence or absence of the ATC alert to the flight crew. This independent variable has three levels in this case study in order to accommodate the examination of the time-to-onset component of the alert:

• Control condition – no wake alert
• With the wake alert – Wake turbulence is imminent (< 1 minute)
• With the wake alert – Wake turbulence is expected in 3 minutes

The resultant hypotheses are as follows:

• H[1] The ATC wake alert will improve crew wake-upset management.
• H[2] The greater improvement in flight crew wake-upset management will be observed when the ATC alert is given at 3 minutes out, rather than when the wake-vortex is imminent.
• H[3] The ATC wake alert will have a smoothing effect on perceived workload: an increase in workload will occur earlier than the wake turbulence

onset, and last longer, but will be lower at its peak than without the ATC wake alert.

- H^4 The ATC wake alert will enhance situational awareness.
- H^5 The ATC wake alert increases the pilot's arousal immediately after the caution, and aircraft piloting will be better (quicker time to respond, less standard deviation in roll, bank angle correction) when responding to the onset of wake turbulence, than without the alert.
- H^6 The ATC wake alert will lower stress, and time to recover to the baseline stress level (transient over time) will be quicker than without the alert.
- H^0 Null Hypothesis: There is no significant difference in flight crew management of the wake upset between having the ATC wake alert and not having the wake alert.

The following sections detail the measures (subjective and objective) selected from the HURID toolkit in order to validate the Wake Alert concept.

Subjective Measures

The Bedford Workload Scale (Roscoe, 1984) elicits subjective ratings of workload from participants by presenting the operator with a 10-element scale. To simplify the process of choosing one of the ten workload ratings, the Bedford scale juxtaposes a hierarchical decision tree onto the ten scale ratings. Operators must navigate through the hierarchy and narrow down their choices of workload ratings to two or three choices and then select a single rating based on the descriptions attached to the ratings. The Bedford Workload Scale will be administered in the debrief, after the end of each run. This scale was selected because of the benefit from the verbal descriptions next to each value and thus it was felt that this better captures relative workload perception. It was not possible to use a workload scale that is administered as a 'probe' during the experimental run, as this would impact on the fidelity of the simulation, moreover a verbal probe would require the flight crew to voice their ratings, therefore possibly influencing each other with their response.

The Situation Awareness Rating Technique (SART: Taylor, 1990) is a post-trial subjective rating technique that uses the following ten dimensions to measure operator SA: Familiarity of the situation, focusing of attention, information quantity, information quality, instability of the situation, concentration of attention, complexity of the situation, variability of the situation, arousal, and spare mental capacity. SART is administered post-trial and involves participants subjectively rating each dimension on a seven-point rating scale (1 = Low, 7 = High) based on their performance of the task under analysis. The ratings are then combined in order to calculate a measure of participant SA. SART was chosen to measure SA as it has been designed and validated for use with Aircrew and hence the concepts in it are recognizable by aircrew. Moreover, it can be administered post-trial as not to interfere with the progression of the simulation event.

A third, study-specific questionnaire is being developed to collect exercise-specific feedback from the flight crews. They will include questions to gather

data on the event itself, but also to gather information to inform the development of the procedures & practices associated with the use of a wake vortex alert, the phraseology of such an alert and the roles/responsibilities involved. Additionally, Subject Matter Experts will observe each run in real-time, in order to give their expert opinion on the difference between the conditions (with or without the alert) according to the pilots' actions. Finally, video recording will be taken of the pilots and the cockpit displays in order to capture communications, behaviours, button presses and other display interactions.

Objective Measures

Electrodermal Activity (EDA) is the electrical measurement of the Skin Conductance (SC). Skin resistance varies with the state of sweat glands, and sweating is controlled by the sympathetic nervous system. If the sympathetic branch of the autonomic nervous system is highly aroused, then sweat gland activity also increases, which in turn increases SC. Continuous decomposition (Benedek et al, 2010) analysis will be applied to estimate the pilots' tonic (SCL) and phasic (SCR) components (Braithwaite et al, 2013; Posada-Quintero et al 2020). SCL is the slow-changing part of the EDA signal, mostly related to the global arousal of the participant, whilst SCR is the fast-changing part of the EDA signal which occurs in relation to single stimuli reactions (Boucsein, 2012). The pilots' EDA signal will be recorded with a sample frequency of 64 (Hz) by the Shimmer3 GSR+ Unit (Shimmer, Ireland) and placing the two electrodes, respectively, on the index and middle finger.

Electroencephalography (EEG) refers to the recording of the brain's spontaneous electrical activity over a period of time. The EEG will be recorded by the Mindtooth system (G.A. 950998) and using water-based electrodes. The brain activity will be collected to assess the pilots' mental states. The EEG data will be firstly pre-processed to remove or correct eventual artifacts, and then the Global Field Power (GFP) will be calculated from the artifact-free EEG dataset to characterise and finally measure the mental states (Borghini et al., 2022; Sciaraffa et al, 2022). Both pilots will wear the EEG and EDA sensors and devices, and the neurophysiological signals recording will be synchronised. EEG and EDA data will estimate synthetic (objective) measures of the pilots' mental workload, stress, and arousal while dealing with the flight simulation under the different experimental conditions, and finally, evaluate the impact of the wake vortex alert. The setup of the sensors and equipment is shown in Figure 4.

The equipment used in this study is easy to wear and reliable, fully compatible with aviation instruments (e.g., headphones, glasses) and has wireless connection (i.e., Bluetooth). The simulator is configured as a B737-800, in which the Captain (Capt.) sits to the left, the First Officer (1st Off.) to the right, with the addition of two observational seats behind the aircrew (Obs1 and Obs2), a Simulator Operator (S/O) seat in the middle behind the crew, and two floor positions, a left-hand side one for collection of neurophysiology data and a right-hand floor position for the collection of eye tracking

Figure 4: The EEG and EDA devices will be synchronized to acquire the pilots' neurophysiological signals while dealing with the flight simulation under the different experimental conditions, and estimate their mental workload, stress, and arousal.

Figure 5: Internal layout of Cockpit Simulator.

data. The S/O position also acts as Air Traffic Control and the Obs1 position also acts as the aircraft cabin staff, for the purpose of communications to/from the pilots (Figure 5).

Eye Movement Tracking

Eye-trackers provide behavioural and psychophysiological measures that can offer insight into pilots visual information process based on eye movement. Specifically, the pilot's fixation or gaze points, dwell times (duration spent fixated on one point), saccade lengths (ballistic eye movement between two points), and scanning entropy (changes to the scan-path patterns) provide insight into visual information processing based on their gaze pattern. The Tobii-2 Eye-tracker used in this experiment consists of head-mounted eye-glasses that direct infra-red illuminations towards the eye, resulting in detectable reflections within the eyes. An infrared camera detects and tracks the pupil center and its movement in relation to the corneal reflection (Muehlethaler and Knetch, 2016), thus calculating the direction of the gaze. Integrated Gyro and Accelerometer sensors allow the eye-tracker to

differentiate between head and eye movements during a dynamic environment. The camera mechanism integrated into the Eye-tracker records the eye movement, pupil dilation and gaze patterns of the pilot and co-pilot with minimal to no distraction or change to their activities. The eye-tracker unit is cleaned, fitted and calibrated to every participant at every crew changeover, to ensure accuracy of the measurements. Two ethernet cables connect the eye-tracker, from the back of the pilots' head, to the laptops that collect the data from the eye-tracker unit. The behavioural data from the eye-tracker, and the psychophysiological data (EEG/EDA) both have annotated markers (with a common time-stamp) that are also synchronized with data from the simulator log, observations, audio and video. The Pilot and Co-pilot's gaze patterns can be used to measure: Wake Upset Management procedures, situation awareness, and work-load conditions. Specifically:

- Wake Upset Management procedures will be evaluated by assessing the monitoring of flight displays (time to first fixation of the Flight Mode Annunciator - FMA - from the onset of the Wake Vortex) and decision-making process (fixation sequence and dwell time of the FMA) during task execution.
- Situation awareness can be evaluated using Endsley's (2015) Situation Awareness Model of three levels of Situation awareness: the perception of elements within a volume of time and space (Level 1); the understanding of their meaning (Level 2); and the projection of the elements' future status (Level 3). The eye-tracker can be used to evaluate the first of these levels, through visual perception and by assessing the allocation of visual attention (Wickens et al., 2008) of the elements of interest (fixation rates and dwell time on the Area of Interest such as the FMA, and scanning entropy).
- Workload can be measured using eye-tracker data related to ocular fixation parameters that change due to the effect of stress on the pilot (Causse et al., 2011; Wickens et al., 2004). The duration, frequency and change in fixation, as well as the number of saccades can be used to measure workload.

CONCLUSION

This paper has illustrated the assessment of human performance via flight cockpit simulations of the design of a new alerting, aiming at reducing the startle effect to Pilots in case of aircraft upset induced by wake turbulence encounters. The evaluation is part of the new SAFEMODE HURID framework, providing a comprehensive and integrated framework for assessing the human factors of new designs in view of informing the designers at every stage of concept development lifecycle. The cockpit flight simulation relies on a high fidelity environment, with type VI Boeing 737-800 simulator facility and cruise simulated wake upset inputs which were calibrated against actual wake encounter flight data. Human Factors measurements include workload, situation awareness, trust, acceptability-based user feedback, as well as psychophysiological measures such as eye tracking, Electro-Dermal Activity

(EDA) and EEG. Combining the analyses of psychophysiological measures, flight parameters, expert observations and subjective pilot feedback, enables evaluation of Flight Crews performance in preparing for, managing or avoiding wake encounter upsets with the new ATC wake alerts, showing the net safety benefits. Early results indicate that the simulations can indeed induce startle effect, and that repeated exposure enables flight crew to overcome startle and manage the situation in a more measured and controlled fashion.

ACKNOWLEDGEMENT

REFERENCES

Benedek, M. and Kaernbach, C. Decomposition of skin conductance data by means of nonnegative deconvolution, *Psychophysiology*, vol. 47, no. 4, pp. 647–658, 2010.

Boucsein, Wolfram. Electrodermal activity. Springer Science & Business Media, 2012.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. Psychophysiology, 49(1), 1017–1034.

Causse, M., Dehais, F., Arexis, M., Pastor, J., 2011. Cognitive aging and flight performances in general aviation pilots. Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development, 18.5, 544–561.

Endsley, M. R. (2015). Final reflections: situation awareness models and measures. *Journal of Cognitive Engineering and Decision Making*, 9(1), 101–111.

Field, J.N., Boland, E.J., van Rooij, J.M., Mohrmann, J.F.W., Smeltink, J.W. (2018) Startle Effect Management. EASA Report EASA_REP_RESEA_2015_3. Cologne: EASA.

Gianluca Borghini, Pietro Aricó, Gianluca Di Flumeri, Vincenzo Ronca, Andrea Giorgi, Nicolina Sciaraffa, Claudio Conca, Simone Stefani, Paola Verde, Angelo Landolfi, Roberto Isabella, Fabio Babiloni (2022). Air Force Pilots' expertise assessment in terms of mental effort requested during unusual attitudes recovery flight training simulation. Safety (under review)

Posada-Quintero, H. F., & Chon, K. H. (2020). Innovations in electrodermal activity data collection and signal processing: A systematic review. Sensors, 20(2), 479.

Rooseleer, F., Schmitz, D. and Kirwan, B. (2021) Human Factors analysis for a new Wake Vortex air traffic alert. Applied Human Factors & Ergonomics Conference (AHFE), in Advances in Safety Management & Human Performance, Arezes, P. and Boring, R. (Eds) July 25-29, pp. 105–113. Springer, Switzerland. https://doi.org/10.1007/978-3-030-80288-2_13

Nicolina Sciaraffa, Gianluca Di Flumeri, Daniele Germano, Andrea Giorgi, Antonello di Florio, Gianluca Borghini, Alessia Vozzi, Vincenzo Ronca, Rodrigo Varga, Marteyn Vangasteren, Fabio Babiloni, Pietro Aricó (2022). Validation of a light EEG-based measure for real-time stress monitoring during realistic driving. Brain Sciences (under review)

Muehlethaler, C., & Knetch, C. (2016) Situation Awareness Training for General Aviation Pilots Using Eye Tracking. IFAC-PapersOnLine, 49(19), 66–71.

Roscoe, A.H (1984) Assessing pilot workload in flight. In AGARD Conference Proceedings Flight Test Techniques, Paris.

Taylor, R.M. (1990). Situational Awareness Rating Technique (SART): The development of a toll for aircrew systems design. In: AGARD Conference Proceedings No 478, Situational Awareness in Aerospace.

Wickens, C., Lee, J., Liu, Y., Gorden-Becker, S., 2004. An Introduction to Human Factors Engineering Second Edition.

Wickens, C. D., McCarley, J. S., Alexander, A. L., Thomas, L. C., Ambinder, M., & Zheng, S. (2008). Attention-situation awareness (A-SA) model of pilot error. *Human performance modeling in aviation*, 213–239.