ANALYSIS OF MULTIMODAL SIGNAL IN ROTORCRAFT COCKPIT

D. Marchesoli[*,◇], M. S. Lukasiewicz[†], C. Talamo[◇], A. Zanoni[◇] P. Masarati[◇]

[◇]Politecnico di Milano - Dipartimento di Scienze e Tecnologie Aerospaziali, Campus Bovisa, Via La Masa 34, 20156 Milano, Italy

[†]Politecnico di Milano - Dipartimento di Meccanica, Campus Bovisa, Via La Masa 1, 20156 Milano, Italy

[*] e-mail: davide.marchesoli@polimi.it

Abstract

A methodology to test and validate the effectiveness of multimodal signals in aircraft cockpits is presented. A set of twelve participants not trained as professional pilots are asked to perform a flight-like tracking task. Then, the effects of visual and aural sensors in acquiring the external information are evaluated through objective and subjective methods. The results are used to draw recommendations to perform a complete analysis of a novel multimodal signal.

1 Introduction

Since the spread of automatic flight control systems, pilot responsibilities have gradually included more and more mission-managing aspects, on top of flying the vehicle, resulting in a consistent increase in the amount of information fed to the aircrew [1]. The need to maintain pilots' Situational Awareness, coupled with Human Centered Design approaches, has pushed avionic designers to introduce Multimodal Signals [2]. Compared to classical instrumentation, Multimodal Signals distribute the information across more senses, stimulating the sight sense, the aural sense and the tactile sense, e.g. through the adoption of active inceptor systems [3].

In the late 1970,s McRuer presented a complete pilot model as an analysis tool for aircraft pilots [4]. Later, Hess took extended the model adding the pilot adaptability to the task [5]. The Hess model is divided into two parts: the central nervous system, responsible for interpreting the signals and setting the regulator gains, and the neuromuscular system, representing the involuntary response of the pilot body subjected to vibrations [6]. Concerning the former, in the original paper by Hess [5], the effect introduced by the human sensors receiving the external signals is represented as a pure latency, accounting for the effect of any sensory channel: visual, aural or tactile. Despite this model being able to handle different cues, only the visual effects have been taken into account, whereas in more modern work also the effects of tactile cues have been investigated [7,8]. The neuromuscular response modeling and its effects on aircraft control have been the subject of numerous works, also by the authors [9–12]. It affects the higher frequency bandwidth of the pilot-vehicle interaction, where the voluntary pilot action is not involved and the purely biomechanical response of the pilot body is in play. While there are indications that suggest that the involuntary response depends on the pilot workload [13,14], which in turn is also a function of the available cueing, this aspect of the pilot response is not of interest for the scope of this paper and it is not further discussed here.

The goal of this work is to develop a procedure for the evaluation of novel cockpit layouts and cueing systems. The responses of the pilot to different stimuli while performing a flight task are compared, to differentiate the processes that interpret the acquired physical information for each signal type. The obtained information is then used to evaluate the effectiveness of the multimodal signal analysis tool for use in the design phase of avionics development. Future developments are foreseen to build an extended pilot model integrating the human response to multimodal stimuli to analyse pilot control strategy, i.e. the value of the crossover frequency, when stimulated by multimodal cues.

2 Human Cognitive Architecture and Workload

The final goal of an innovative cueing system is to reduce the overall workload to which the pilots are subjected and, consequently, to increase the performances. Thus, the measure of the workload appears a natural way to evaluate the effectiveness of the signal and its acceptance by pilots, in case the workload is also evaluated using subjective scales. Since no universally accepted definition of the workload exists [15], here the workload is defined as:

> *the degree of saturation of the working memory, limited in capacity, in order to accomplish a given task subjected to the performer's internal characteristics, environmental and situational factors.*

This is a reduced and adapted version of the inclusive definition contained in [15], rearranged specifically for this task.

The cognitive load can be categorized as intrinsic or extraneous, as stated in the Cognitive Load Theory of Sweller [16]. The intrinsic cognitive load refers to the complexity of the task being carried out without any reference to how the task is performed. Thus, by its own nature, the intrinsic cognitive load is fixed once the task and the ability of the performer are fixed. The more the elements of a task interact with themselves, the more the task increases in complexity and the intrinsic cognitive load is high. On the other hand, the extraneous cognitive load is something representing how the task is being carried out, how the performer has been instructed and, the main interest of this work, how the performer is cued of their performance. Multiple effects affect the amount of extraneous cognitive load (all the details may be found in [16]), but for this specific research only a few are of interest:

- the split-attention effect: when exposed to two or more sources of information unintelligible in isolation, the performer has to integrate them to carry out the task increasing the extraneous cognitive load;
- the modality effect: the information fed can be presented in different modalities, for example visually and aurally; the multimodal presentation of the information should increase the available working memory and, consequently, reduce the total workload;
- the redundancy effect: this effect is the dual of the modality effect; if the information is unnecessarily fed in different modalities, the interaction between the elements is increased, thus increasing the total workload;
- the expertise reversal effect: the level of expertise of the performer can affect the effectiveness of a cueing system since the inclusion of a source of information that may be fundamental for non-experts can be detrimental for highly skilled performers because a useless amount of information must be processed and discarded.

In the analysis of a multimodal signal, all these effects must be taken into account, avoiding the undesirable result of increasing the workload of the pilot instead of reducing it.

Beyond the tasks and the effects induced by a cueing system, the allocable working memory is affected by other internal characteristics of the operator, such as emotional status. Violent, sudden, and unexpected emotions caused by particular events may overwhelm even an experienced pilot, leading to unforeseeable reactions. A clear example is the accident that occurred on Air France Flight AF 447 [17], during the cruise phase the crew encountered a failure of speed measurement and a consequent autopilot disengagement resulting in a high workload situation in which the flight control law reversed to the Direct Mode. The inability of the crew to immediately identify the problem led to a violent emotional reaction which corresponds "to a threshold in terms of being able to take into account an unusual aural warning. In an aural environment that was already saturated by the C-chord warning [..]". This situation can be explained also referring to the Cognitive Tunnel Effect, in which, under heavy stress, the attention is fully focused on visual cues [18].

3 Experimental Approach

The experiments are aimed at distinguishing the effectiveness of the cueing signals presented to the pilot for a tracking task, depending on the selected sensory channel.

The Mission Task element (MTE) consists of pitch attitude tracking using a linearized helicopter model, derived from [19], maintaining the Stability Augmentation System active, such that the attitude tracking can be performed also by people not trained as pilots, controlling the helicopter attitude, i.e., using an Attitude Command Attitude Hold (ACAH) setting. The MTE is performed in forward flight starting at 60 knots airspeed and is kept simple enough to be flown not only by professional pilots but also by people not trained as pilots – also referred to as non-pilots in the following. The participant controls only pitch and roll using the cyclic stick; the other inceptors' signals are fixed in the trim position.

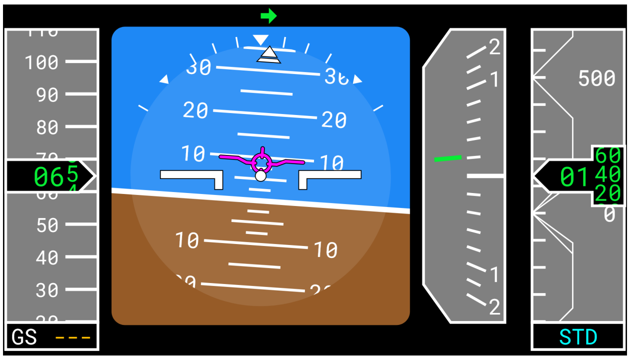

Figure 1: Screen capture of the display used, with the visual cue for target pitch shown as the magenta marker in the Attitude Indicator.

Due to the nature of the experiment, having received professional pilot training is not essential and is expected to affect the results of signal perception to a limited extent only. Moreover, the participation of non-pilots is useful mainly for two reasons: (1) non-pilots are not used to classical instrumentation, thus they show no preference or bias for any signal; (2) non-pilots availability is higher and therefore it is easier to involve a greater number of tested subjects, making the results statistically more meaningful.

3.1 Pilot cues

The experiments were conducted on a fixed base simulator using classical helicopter control inceptors, as shown in Fig. 2. The signals were primarily given to the pilot using an open-source visual display[1] resembling the classical instrumentation of a glass cockpit. In all trials, the participants were shown the attitude indicator along with other instruments they did not need to actively monitor: airspeed, vertical speed and altitude indicators. Depending on the tested configuration, the target pitch was communicated using a visual cue, an auditory cue, or both. The visual cue involved a

[1] Python package lidia: `https://pypi.org/project/lidia/`

Figure 2: A photo of one of the participants performing the experiment with both visual and audio cues.

magenta-colored marker presented in the attitude indicator, which had to be tracked using the pitch-roll symbol (Fig. 1).

The auditory cue was designed to mimic the sound produced by variometers, or vertical speed indicators, used for example in paragliding, which provides continuous audio feedback to facilitate finding thermal currents. This is achieved by modulating the sound frequency and interrupting the produced sound with a fast on-off rhythm, where a higher pitch and faster rhythm indicate a faster ascent. To use this cue as feedback for pitch tracking, when the pitch was higher (i.e., nose-up) than required, the produced sound was similar to an ascending glider. The real-time sound synthesis based on the simulation state was done with Unreal Engine 5 software.

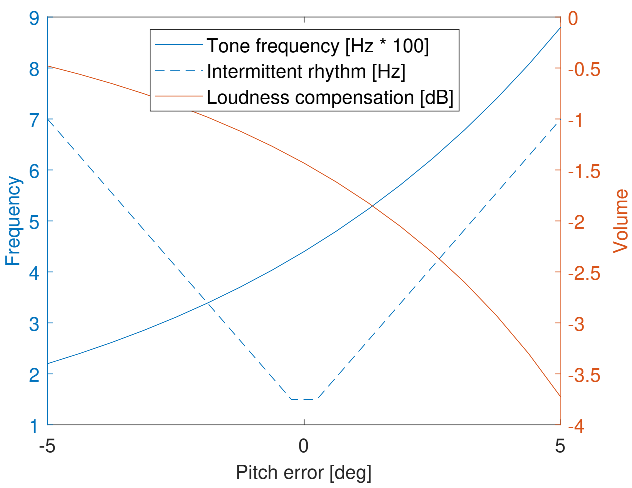

Figure 3: Plot of all parameters of audio synthesis in the feedback range. Values outside this range were extrapolated as constant.

Based on task performance and participant surveying in the pre-trials, the audio cue was modified from the sound of the variometer. The main issue was high frustration and a sense of urgency specifically when the aircraft was piloted with the pitch above the target attitude, producing a high-pitched sound. To limit the fatigue caused by continuous exposure, the range of tones was lowered to fit between the A_3 $(220\,\mathrm{Hz})$ and A_5 $(880\,\mathrm{Hz})$ musical notes. Additionally, a volume curve was applied to compensate for the difference in perceived loudness of different frequencies. Unlike gliding, where a slow descent is a normal state, the tracking task was symmetrical with respect to the error direction, therefore the rhythm was changed to increase with error magnitude both for positive and negative errors. The final parameters used in the analysis are shown in Fig. 3.

3.2 Experimental procedure

To provide a signal at the required frequency content for system identification, the target pitch was generated as a sum of multiple sine waves, that had a random phase, and their amplitude defined with a power spectral density (PSD) function. The power in the signal was defined to be unity between the frequencies $0.05\,\mathrm{Hz}$ and $0.2\,\mathrm{Hz}$, then decreasing linearly with frequency until reaching zero for $0.4\,\mathrm{Hz}$. To construct the signal, a sine wave was added for every component frequency in the PSD domain. To determine the amplitude, the PSD function was integrated between each component frequency for this specific signal component, and then the sum of all of them was scaled to provide target excursions up to $\pm 6^\circ$.

The experiment involved 12 participants (3 female), recruited among university students and staff. All of them had normal hearing and normal or corrected to normal vision. One of the participants, marked with "(H)" in the figures, is the holder of a private pilot license for helicopters, while the other participants had no formal piloting training. Participation in the study was voluntary, with no compensation provided.

The experiment procedure started with ian explanation of the displayed visuals, and a training run to accustom the participant to helicopter dynamics and the tracking MTE. During all of the training runs, both visual and auditory cues were available. All participants were instructed to continue the training run until they felt that they had learned the task and were not improving anymore; however, a minimal training time was set to $120\,\mathrm{s}$.

After the initial training, each participant performed a recorded trial for a duration of $120\,\mathrm{s}$ in each of the following configurations:

1. visual cue
2. visual and audio cue
3. audio cue

To balance factors dependent on trial order, e.g., training and fatigue, the order of the configurations for any given participant was randomly selected without repetition from 6 permutations of three-element set. With a total of 12 participants, each permutation was performed twice.

4 Data Analysis

The simulation was run using a state-space model in Simulink at a fixed rate of $1000\,\mathrm{Hz}$. The model's attitude angles in degrees and dimensionless controls deflection δ between -1 and 1 were saved at every timestep. After every trial an NASA TLX questionnaire[2] was completed for the specific configuration tested. Additionally after completing the series of three trials, each participant was asked an open question to state their preference and other comments on the provided cueing modes.

The collected data are analysed using both objective and subjective methods. The objective methods are necessary to assess the effectiveness of the cue independently from the operator's opinion and emotional status. Moreover, the objective evaluation of the signal is fundamental as the volunteers used for the trial were not formed as Flight Test Pilots and, thus, received no specific formation about how to auto-evaluate their performance using subjective scales. The last consideration can be also applied to professional pilots holding a commercial ATPL licence without any formation as Flight Test Pilots and must be taken into account considering the Expertise Reversal Effect: the cues tested with highly skilled pilots, as FTPs are, may be judged differently from what commercial pilots, or even persons without any formation as pilots, would. Objective methods include:

- the Root Mean Square (RMS) error: the RMS error is useful to evaluate the overall performance;
- the Frequency Response Function: the FRF is calculated using the error as input and the operator action as output, as in Fig. 4;
- the pilot Aggressiveness: it is calculated as $\frac{1}{T}\int_T |\dot{\delta}|\,dt$, where T is the period of time over which the aggressiveness is calculated and $\dot{\delta}$ is the control deflection time derivative.

Despite the importance of the objective methods, the subjective ones should not be discarded as they return the important feedback of pilot acceptance of the proposed cues. Again referring to the AF447 accident final report [17], the investigators clearly state that "During previous events studied, crews frequently mentioned their doubts regarding the relevance of the stall warning", as if a good number of pilots have not accepted the stall warning cue and its implications. The subjective scale used is the NASA TLX because it is specifically designed for generic tasks and not only for flight tasks as the Bedford scale. Moreover, also the personal opinions and considerations of the performers have been registered as an insight into their preferences and acceptance regarding the cues.

[2]Implementation used: `https://www.keithv.com/software/nasatlx/nasatlx.html`

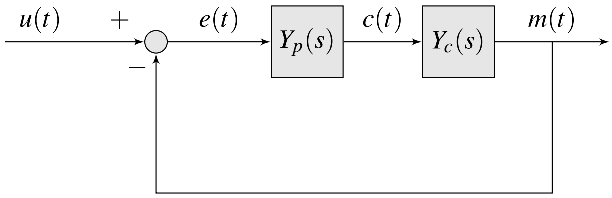

Figure 4: The block diagram representing a single-loop compensatory human control. The error, $e(t)$, processed by the human operator is the difference between the reference signal, $u(t)$, and the output, $m(t)$. The pilot, $Y_p(s)$ compensates for the error feeding the controlled signal $c(t)$ to the controlled element transfer function, $Y_c(s)$.

5 Results

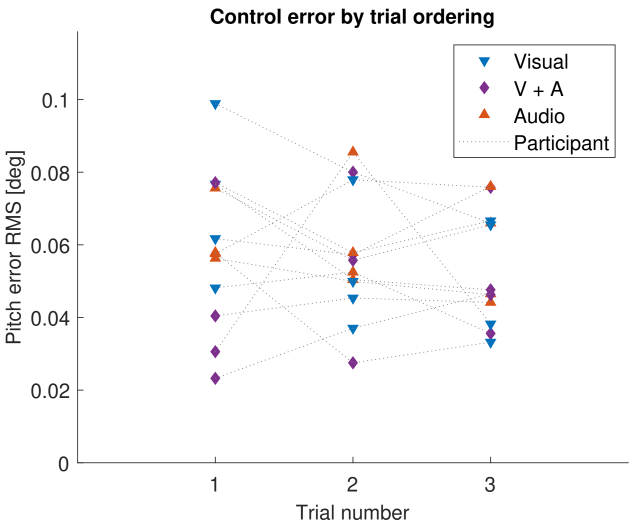

Figure 5: Tracking error RMS considering the trial order.

As anticipated in the Experimental Procedure section 3.2, the order of the single task is permuted to check and avoid possible training effects. Referring to Figure 5, we can conclude that the training effects may be discarded for this specific trial since no clear trend can be identified. The aggressiveness on the controls shows no particular trend, ref. Figure 6, and it is consistent among all participants, with the exception of operator P01 who is particularly aggressive in all trials. It is of great interest the behaviour of P01 with the aural-only cue; the excessive aggressiveness the subject is exerting, without any reduction in the control error (Fig. 7), may be explained as a "wait and act" strategy: the pilot waits for the error signal and then acts very aggressively on the control. On the other hand, the RMS error (Fig. 7) is consistent among all the participants up to an error threshold, more or less around $0.5°$. Below this error threshold, the aural-only cue returns the highest error, the visual-aural cue the lowest and the visual-only one is very close to the visual-aural cue. Above the threshold, the RMS error trend shows an inversion: the aural-only cue returns the lowest error and the visual-only cue the highest one. A similar outcome has been detected also in other work about

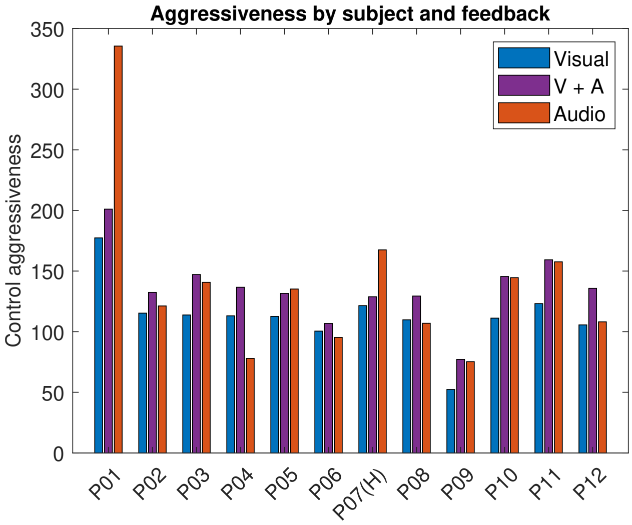

Figure 6: Pilot aggressiveness on controls for each subject.

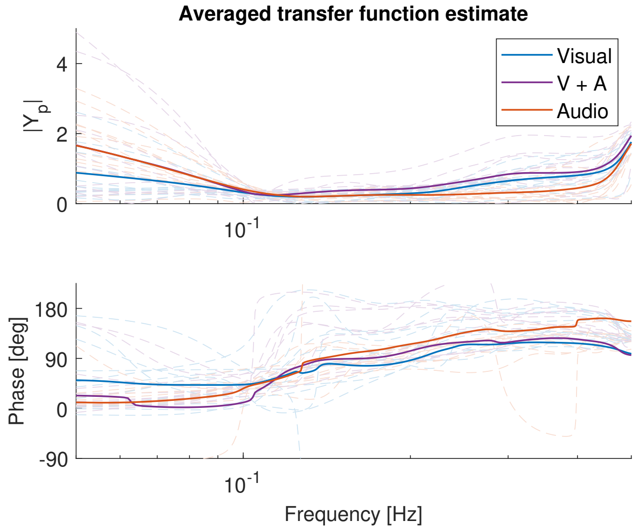

Figure 8: Average frequency response function using the error, $e(t)$, as input and the controlled element, $c(t)$, as output.

multimodal signals [20]; it can be explained by how the error is detected by the operator: with the aural-only cue, the error is indirectly reconstructed by the operators themselves, whereas the visual cue, with or without the aural one, returns immediate error feedback. Some subjects may react to the error visualization by feeling excessive stress affecting their performance.

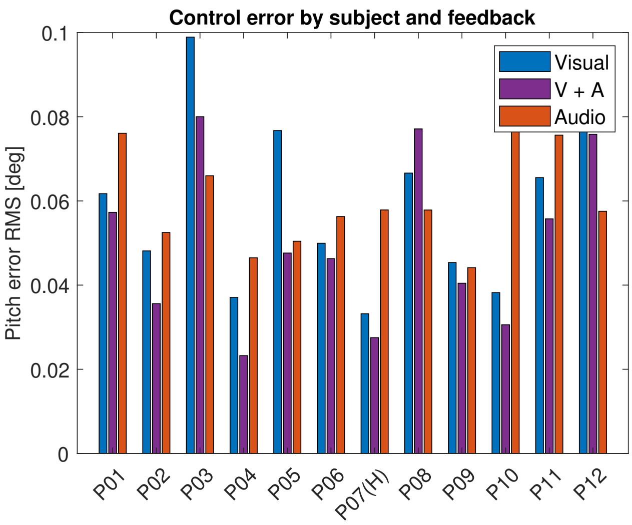

Figure 7: RMS error for each subject.

The frequency response function has been calculated for each subject; for the sake of clarity, the average for each cue type is highlighted in Figure 8. The results are consistent among all participants. The visual and visual-aural cues show similar results with an increase in the error at the higher frequency of the target bandwidth. The apparent decrease of the error at high frequency with the aural-only cue is due to the lack of action from the pilot. The aural cue appears to act as a low-pass filter on the pilot action.

The trend of the NASA TLX scores, in Fig. 9, is more consistent among all the tested subjects but it does not confirm any of the previous considerations on the objective indices. This inconsistency can be explained by the little familiarity the tested subjects had with the scale and with the self-evaluation of a task. On the other hand, it can be also a symptom of the pilot acceptance of the cue, with lower acceptance of the aural one, as confirmed by open-question responses.

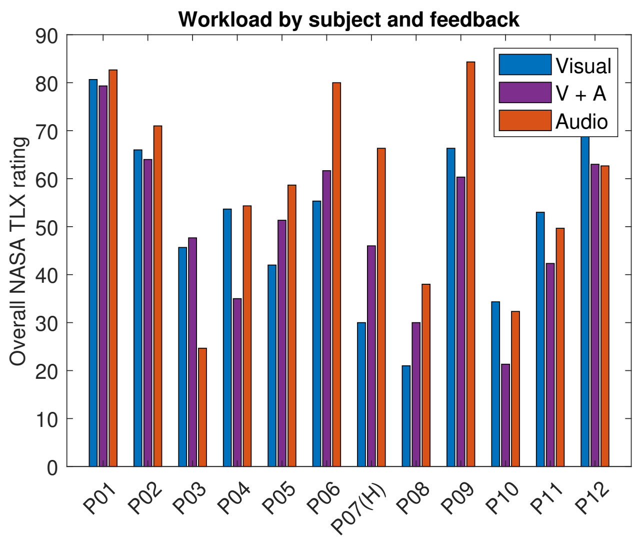

Figure 9: NASA TLX outcomes for each subject.

6 Conclusions

Despite the limitations of this work:

- small amount of processed data;
- insufficient number of tested subjects for a statistically significant analysis;

- no professional pilots tested;
- single tracking task with only aural and visual cues;

some preliminary conclusions and recommendations may be drawn about the analysis technique of a multimodal signal:

- pilot emotions and acceptance of the cue may heavily affect the results;
- cues may act as a filter on pilot control;
- aggressiveness and precision are inversely proportional only if the pilot does not apply a "wait and act" strategy.

The analysis technique needs to be expanded, using other tasks and cues, to be formalized in a standard procedure.

References

[1] Anderson, E., Fannin, T., and Nelson, B., "Levels of Aviation Autonomy," 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC), 2018.
doi: 10.1109/DASC.2018.8569280

[2] Smith, C. J., "Design of the Eurofighter human-machine interface," Cockpit Displays V: Displays for Defense Applications, edited by D. G. Hopper, Vol. 3363, 1998.
doi: 10.1117/12.321768

[3] Barbé, J., Spaggiari, L., Clay, A., Bérard, P., Aissani, A., and Mollard, R., "Why and how to study multimodal interaction in cockpit design," Proceedings of the 15th Ergo'IA" Ergonomie Et Informatique Avancée" Conference, 2016.

[4] McRuer, D. and Krendel, E., "Mathematical Models of Human Pilot Behavior," Technical Report AD-775905, NATO Advisory Group For Aerospace Research And Development, Paris, France, January 1974.

[5] Hess, R. A., "Structural Model of the Adaptive Human Pilot," *Journal of Guidance and Control*, Vol. 3, (5), 1980, pp. 416–423.
doi: 10.2514/3.56015

[6] Venrooij, J., Abbink, D. A., Mulder, M., Van Paassen, M. M., Mulder, M., Van Der Helm, F. C. T., and Bulthoff, H. H., "A biodynamic feedthrough model based on neuromuscular principles," *IEEE Transactions on Cybernetics*, Vol. 44, (7), Cited by: 11, 2014, pp. 1141 – 1154.
doi: 10.1109/TCYB.2013.2280028

[7] Chaudhari, R., Steinbach, E., and Hirche, S., "Towards an objective quality evaluation framework for haptic data reduction," 2011 IEEE World Haptics Conference, 2011.
doi: 10.1109/WHC.2011.5945543

[8] Efremov, A., Aleksandrov, V., Efremov, E., and Vukolov, M., "The influence of different types of inceptors and their characteristics on the pilot-aircraft system," *IFAC-PapersOnLine*, Vol. 51, (34), 2nd IFAC Conference on Cyber-Physical and Human Systems CPHS 2018, 2019, pp. 372–377.
doi: https://doi.org/10.1016/j.ifacol.2019.01.013

[9] Zanoni, A. and Muscarello, V., "Moving towards A-Priori Identification of Undesirable Pilot Biometrics for Collective Bounce Instability," 43rd European Rotorcraft Forum, ERF 2017, Vol. 1, 2017.

[10] Zanoni, A., Cocco, A., and Masarati, P., "Numerical investigation on the role of control inceptors layout in rotorcraft-pilot couplings," 77th Annual Vertical Flight Society Forum and Technology Display, FORUM 2021: The Future of Vertical Flight, Cited by: 1, 2021.

[11] Zanoni, A., Cocco, A., Marchesoli, D., Talamo, C., Masarati, P., Colombo, F., Kemp, S., and Fosco, E., "Pilot Biomechanics for the Definition of a Rotorcraft-Pilot Interaction Experiment," 78th Vertical Flight Society Annual Forum and Technology Display, FORUM 2022, Cited by: 0, 2022.

[12] Zanoni, A., , Marchesoli, D., Talamo, C., Masarati, P., Colombo, F., Kemp, S., and Fosco, E., "Experimental Test-Bed for the Identification of Biodynamic Feedthrough of Helicopter-Pilot Systems," 48th European Rotorcraft Forum, Winterthur, Switzerland, 6-8 September, 2022, 2022.

[13] Venrooij, J., Abbink, D. A., Mulder, M., van Paassen, M. M., and Mulder, M., "Biodynamic feedthrough is task dependent," 2010 IEEE International Conference on Systems, Man and Cybernetics, October 2010.
doi: 10.1109/ICSMC.2010.5641915

[14] Zanoni, A., Zago, M., Paolini, R., Quaranta, G., Galli, M., and Masarati, P., "On Task Dependence of Helicopter Pilot Biodynamic Feedthrough and Neuromuscular Admittance: An Experimental and Numerical Study," *IEEE Transactions on Human-Machine Systems*, Vol. 51, (5), October 2021, pp. 421–431.
doi: 10.1109/THMS.2020.3044971

[15] Longo, L., Wickens, C. D., Hancock, G., and Hancock, P. A., "Human Mental Workload: A Survey and a Novel Inclusive Definition," *Frontiers in psychology*, Vol. 13, 2022, pp. 883321.

[16] Sweller, J., "CHAPTER TWO - Cognitive Load Theory," *Psychology of Learning and Motivation*, edited by J. P. Mestre and B. H. Ross, Vol. 55, Academic Press, 2011, pp. 37–76.
doi: https://doi.org/10.1016/B978-0-12-387691-1.00002-8

[17] BEA, "On the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF 447 Rio de Janeiro - Paris," Technical report, 2012.

[18] Dirkin, G. R., "Cognitive Tunneling: Use of Visual Information under Stress," *Perceptual and Motor Skills*, Vol. 56, (1), PMID: 6844068, 1983, pp. 191–198.
doi: 10.2466/pms.1983.56.1.191

[19] Padfield, G. D., *Helicopter Flight Dynamics: Including a Treatment of Tiltrotor Aircraft*, John Wiley & Sons, third edition, November 2018, pp. 278–282.

[20] Prizio, F. W., Lukasiewicz, M. S., Xia, Q., Garbo, P., and Quaranta, G., "Haptic Device Embedded in Rotorcraft Seats for Target Acquisition and Tracking Tasks," Aerospace Europe Conference 2023 - 10th EUCASS - 9th CEAS, July 2023.