33RD CONGRESS
OF THE INTERNATIONAL COUNCIL
OF THE AERONAUTICAL SCIENCES
STOCKHOLM, SWEDEN, 4-9 SEPTEMBER, 2022

ICAS
2022
SWEDEN

EXPLORATORY STUDY OF MENTAL WORKLOAD IN HELICOPTER AIR-TO-GROUND ROCKET FIRING

Raphael Gomes Cortes[1,2], Emília Villani[1] & Moacyr Machado Cardoso Júnior[1]

[1]Aeronautics Institute of Technology(ITA), São José dos Campos/SP – Brazil
[2]Master Degree Candidate in Aeronautical and Mechanical Engineering Program

Abstract

Helicopter rocket-firing requires continuous pilot training for mission accomplishment purposes, so measuring the mental workload involved could help to improve training effectiveness and flight safety. Physiological and flight data from two Brazilian Army pilots with distinct experience levels have shown that the aiming phase is the more mental demand, and some physiological patterns were identified. Pearson's Correlation and Principal Component analysis, including shot stability parameters (aiming path area, perimeter, and deviation of maneuver parameters) and physiological data, have identified electrodermal activation as the more consistent human dimension related to the shooting performance.

Keywords: Firing, Mental Workload, Physiological sensors, Rocket

1. Introduction

Military flight maneuvers are known to be riskier and rely on individual performance. Indeed, pilots must do continuous training to acquire the required cognitive skill levels. Studies ([6], [8], [12], [16], and [20]) have pointed out that pilot´s mental workload could be differentiated between mission segments and between pilots with distinct levels of skills. Mental workload measurements in military aviation are important for improving flight safety, reviewing doctrine, and validating the feasibility of training in a flight simulator [13]. High levels of cognitive demand can lead to errors with catastrophic outcomes. Thus this knowledge can help avoid errors [20].

A way to understand this training process and the required experience level is to measure the mental workload of these pilots using qualitative and quantitative techniques, thereby contributing to extracting relevant cognitive patterns of correlation in pilot decision making.

In this study, helicopter air-to-ground rocket-firing maneuver performed by the Brazilian Army was recorded during a training section. The pilots must navigate the aircraft towards the target and acquire a certain height, airspeed, distance from the target, and attitude; this requires cockpit coordination and mental workload. Physiological data were collected from two experienced pilots who each performed a sequence of operational maneuvers with diverse levels of engagement.

As piloting is a complex and dynamic environment, several quantitative measures were collected using eye tracker sensors, electrocardiogram (ECG), respiration rate (RESP), electrodermal activity (EDA), and qualitative data was collected using the NASA-TLX scale. These sensors measure the balance of the sympathetic and parasympathetic autonomic nervous systems. Aside from mental workload measurements, recorded flight data was also available and used for obtaining pilot performance on the stabilization of the aircraft during the shot.

The mental workload was assessed through mission segment and was investigated relations of performance on the shot and physiological measures through Pearson's correlation coefficient and Principal Component Analysis (PCA).

To assess pilot performance on a helicopter, rocket firing was used as reference rifleman stability shot parameters ([5] [9]). In this, the shot stability is measured as the area of the aiming points on the target and was also analyzed by the perimeter of these points. It is known qualitatively by rifle experts that breath control, firing during the natural respiratory pause, is a fundamental shot parameter for a good shot [9]. Another parameter analyzed was the deviation of the maneuver parameters [12].

2. Materials and methods

2.1 Participants

Two experienced helicopter pilots from the Brazilian Army volunteered to participate in the physiological measurements during a routine training section. A summary of the demographic data of these participants is presented in Table 1. Both are operational combat pilots, and as could be noted, they have similar flight experience. Still, one pilot had his last rocket shot ten years ago, and the other had done a maneuver practice just the day before.

Table 1 – Summary of pilot's experience.

Pilot	Age	Flight hours	Last shot
1	38	1100	10 years
2	35	700	1 day

2.2 Scenario

The measurement of physiological signals occurred during a Brazilian Army Aviation helicopter air-to-ground rocket firing exercise. The flight section starts with a series of training shots, and in the second phase, after learning the maneuver basis, a series of operational rocket fire maneuvers were performed. The pilot performs a low-level flight using the ground reference to navigate on that maneuver. Then, close to the mission target, he rapidly ascends the aircraft to aim at the target and fire rockets. To finalize, he performs an evade curve. Therefore, each maneuver, as shown in Figure 1, could be divided into five phases: level flight approximation phase, Final leg at low-level flight, climb, aiming and shot stage, and evasion phase. During the aiming phase, the pilot must seek previously established values of airspeed (v), height above ground (h), and distance (d), which are maneuvers parameters, and adjust horizontal attitude towards the target, compensating positioning errors, seconds prior the shot event. During the climb, some G-loads are reached, less than 1.5 G's, positive load when the aircraft ascends from low-level flight to shot height and negative on the reversion aircraft's attitude when starting the aiming phase.

Each pilot performed a sequence of training trials and executed three operational maneuvers with diverse levels of engagement and fatigue at the end of the flight. For analysis, the training maneuvers were excluded.

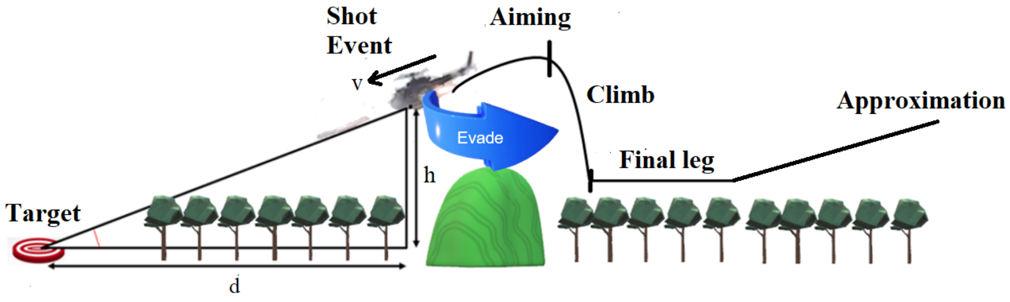


Figure 1 – Maneuver scenario profile.

2.3 Apparatus

The aircraft used in rocket firing training was the AS550A2 Fennec AvEx model in the armed configuration (Figure 2), with a projected aiming sighting system. This aircraft was dual pilot operated with stability augmented systems working to compensate helicopter instability in pitch and roll movements. Flight Data was recorded from the aircraft system to post-flight analysis with a sampling rate of

4 Hz. In this study, the sensors used were Tobii Pro Glasses 2 for ocular measures and TEA-CAPTIV sensors for physiological data (electrocardiogram (ECG), respiration rate (RESP), and electrodermal activity (EDA)). The sensors allocation representation inside the cockpit is presented in Figure 3.

Figure 2 – AS550A2 Fennec AvEx at rocket firing configuration.

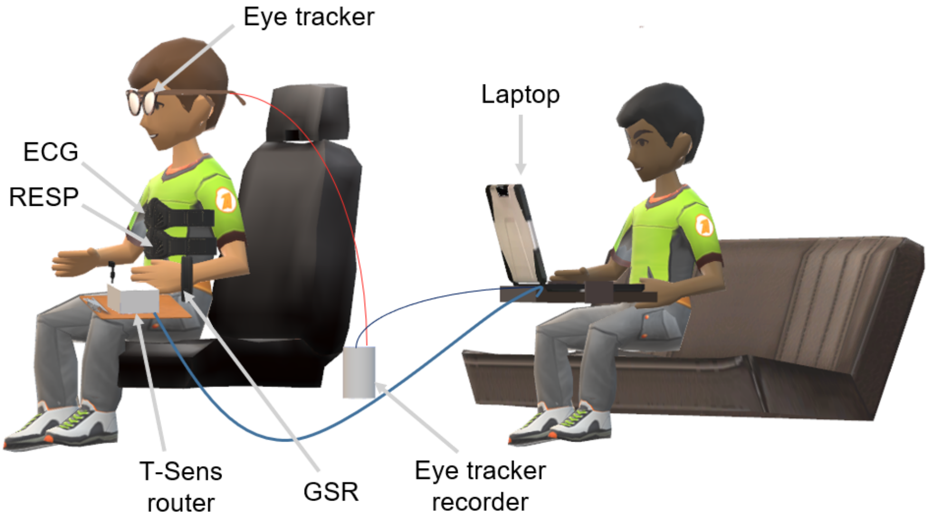


Figure 3 – Sensors allocation.

2.4 Procedures

Initially, the participant was briefed about the research purpose and the physiological measurements; then, he formally expressed his voluntariness. The participant was asked to fill in demographic data in the sequence, and the NASA-TLX scale was explained. In this training section, the participant and the researcher should enter the aircraft with the engine running after the training of another pilot ("hot seat") because of weight and balance issues; hence all the sensors were placed on the ground. As the boarding occurred, the researcher did the calibration and started the recording, while rockets ammunition loading happened before the departure. The researcher did not intervene during the flight, and events were marked through TEA-CAPTIV software. In this sense, the NASA-TLX scale was filled just after the end of the flight.

2.5 Data Analysis

2.5.1 Performance

The flight data recorder was used to estimate the helicopter aiming point on a reference plane at ground level and the impact point on the real target. This was achieved through trigonometric formulas based on latitude (Lat), longitude (Lon), height above ground (h), heading (ψ), and aircraft attitude (θ) as indicated in Figure 4.

As the purpose was to assess aircraft shot stability, the adopted simplification hypothesis was that the rocket trajectory hit the line of sight, and no wind correction was used in calculations.
Based on aircraft position (latitude and longitude) was determined aiming point distance on the reference plane by 1 and these measures were put onto a digital map, converting these distances from the aircraft into GPS coordinates applying the Haversine formula.

$$d_{reference\ Plan} = h \cdot tan(90^o - \theta) \tag{1}$$

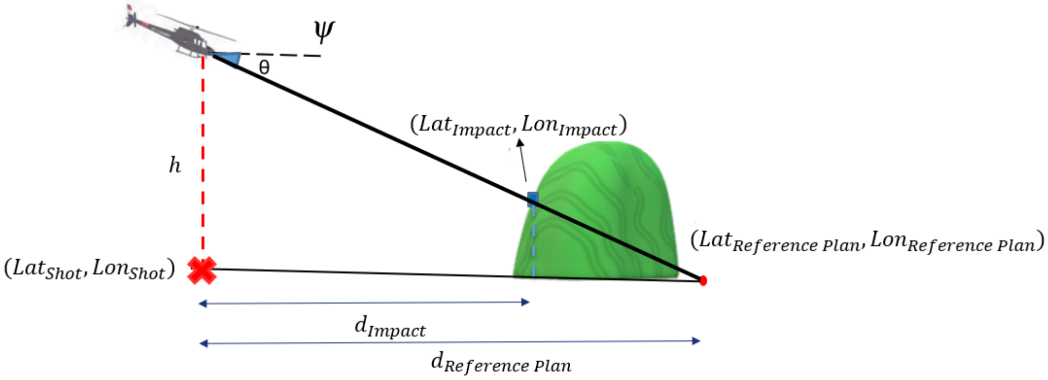


Figure 4 – Aiming point estimation.

The comparison of the real impact point, verified on eye tracker video, and the rocket fired estimated impact point was proposed for validation purpose of the estimation.
To measure the overall steadiness of each shot aiming was calculated as the area of the aiming holding box (Figure 5) and the perimeter from the set of points two seconds before trigger pull, as is used for rifleman shot ([9], [5]). Another calculated flight performance score was the deviation from maneuver aiming parameters (F) at shot moment [12]. Equation 2 determines F-score (F) using airspeed $\overline{\Delta v}$, height above ground $\overline{\Delta h}$, and distance $\overline{\Delta d}$, where $\overline{\Delta X}$ is the normalized maximum deviation expected given by 3. The higher the score, the better the pilot's performance in acquiring maneuver parameters.

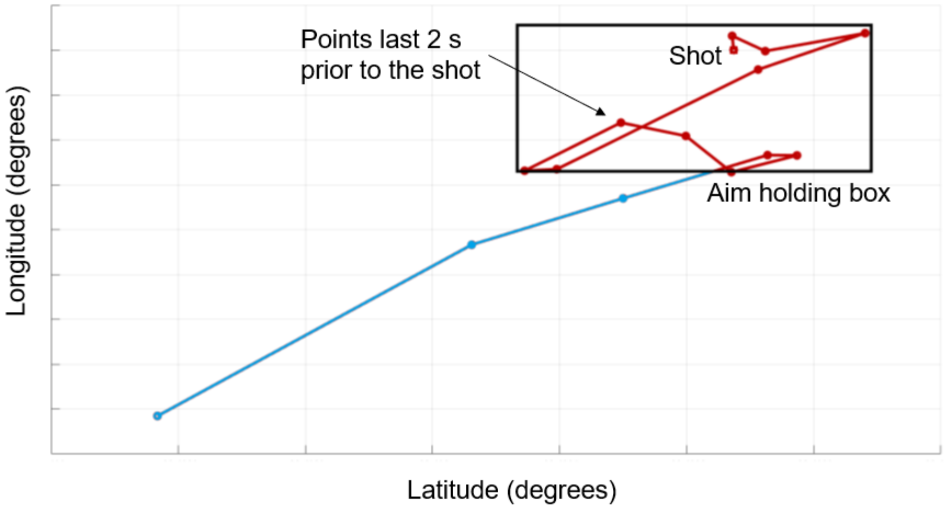


Figure 5 – Aiming holding box.

$$F = 3 - (\overline{\Delta v} + \overline{\Delta h} + \overline{\Delta d}) \tag{2}$$

$$\overline{\Delta X} = \frac{X\ Deviation\ from\ maneuvre\ reference}{X\ Maximum\ acceptable\ deviation} \tag{3}$$

2.5.2 Data standardization

Most physiological parameters have a strong individual component, which means that absolute values can hardly be compared between subjects [2]. This requires transforming the data into a standard scale to allow a comparison between subjects. One of the common performing methods is the Z-transformation, shown in 4, where x_i is the original data, $\bar{x}$ is the individual's average, s is the individual's standard deviation, μ is the group average and σ is the group pooled standard deviation. This formula transforms the data so that all participants have the same average values and standard deviation as the whole group [13].

$$Z = \frac{x_i - \bar{x}}{s}\sigma + \mu \tag{4}$$

The standardization was performed for the overall maneuver trial sixty seconds before the shot and twenty seconds after to reach the approximation phase to the evade curve phase.

2.5.3 Statistical analysis

The data Z-transformed of the two pilots was gathered to evaluate patterns over flight mission segments. Statistical analysis was performed using ANOVA after check residuals assumptions of randomness, homoscedasticity, and normality were met. Also effect size was included in the results as partial eta squared (η_g^2) (conventions: Small (S): $d < 0.01$, Medium (M): $0.06 < d < 0.14$, Larger (L): $d > 0.14$) and Cohen´s d for pairwise comparisons (conventions: Small (S): $d < 0.5$, Medium (M): $0.5 < d < 0.8$, Larger (L): $d > 0.8$).

The performance and physiological relations at the aiming phase were verified using Pearson´s correlation coefficient and Principal Component Analysis (PCA) using data Z-transformed.

2.5.4 Respiration activity

The premeditated respiratory frequency rate $freq(t)$ [16] was the respiratory feature assessment of rocket fire maneuver because respiration patterns could be modified by speech and motion [14]. Hence to compute this, respiration data (RESP) were band-pass filtered (0.05–0.50 Hz, zero-phase), then was applied the continuous wavelet transform (CWT), using the analytic Morse wavelet, which gives the distribution of signal power as a function of time. CWT computation was done directly from MATLAB's wavelet functions. The mean power of each frequency band was normalized by the overall spectral power at that time [16], as shown in 5.

$$freq(t) = \frac{\sum_{i=0.05}^{0.5} Spectral\ Power_{i,t} \cdot freq_i}{\sum_{i=0.05}^{0.5} Spectral\ Power_{i,t}} \tag{5}$$

2.5.5 Electro cardiac activity

Heart rate variability (HRV) analysis of the electro cardiac activity was performed using MATLAB's toolbox PhysioNet Cardiovascular Signal [19], where R peaks were identified and converted into NN signal.

The HRV time domain is usually applied for long-term data (i.e., 24 h or longer) or short-term (i.e., 5 min) recording [19], which both are prohibitive to the application of rocket shot maneuver because the time difference between two shot trials is less than five minutes. For this reason, ECG time domain parameters (mean heart rate frequency (HR), root mean square deviation (RMSSD), and the standard deviation of normal to normal (SDNN)) were computed by an ultra-short-term heart variability analysis using a sliding window of sixty seconds, with one second step, which has been proven to be the minimum reliable ultra-short-term heart for time-domain HRV features [18]. The sixty seconds sliding window ($w(t)$) can be denoted as indicated in 6, where x(t) is the HRV time-domain feature at time t.

$$y(t) = \sum_{i=-30}^{30} x(t)w(t+i) \tag{6}$$

Poincaré features (SD1: standard deviation of instantaneous beat-to-beat variability, SD2: long-term standard deviation of continuous RR intervals) were also computed using the same methodology, as is expected strong correlation with time domain features [17].

HRV frequency domain parameters assessed were the power spectral density over the Low frequency (LF) band (0.04 to 0.15 Hz), High frequency (HF) band (0.15 to 0.5 Hz), the ratio between low and high-frequency bands (LF/HF) ([1] and [11]) and middle frequency (MF) band (0.07 to 0.14 Hz)[6].
Hence the uneven NN signal was re-sampled at a 7 Hz rate [7] and then applied the Short-time Fourier transform (STFT), which performs Fourier Transform over overlapped windowed segments. STFT window was chosen based on the uncertainty principle, stated in 7, from which was calculated a window time resolution (Δt) of 12.5 s (88 samples), considering 0.04 Hz of frequency resolution (Δf), which is lower than the lower limit of the LF band [1].

$$\Delta t \Delta f \geq \frac{1}{2} \tag{7}$$

2.5.6 Electrodermal activity

Electrodermal activity (EDA), regarded as a sensitive and valid indicator to detect variations in arousal changes [4], can be broken down further into different components: Skin conductance level (SCL) (tonic component) related to the slower characteristics of the signal and Skin conductance response (SCR) (phasic driver) refereed to faster changing elements of the signal [6]. Only the SCR component was analyzed to assess short time maneuver mission segments in this study.
EDA data were analyzed using MATLAB's software LEDALAB [3]. Initially, the data was filtered with a first-order Butterworth low pass filter with a cut-off frequency of 5 Hz, and artifacts were manually removed. SCR was extracted through Continuous Deconvolution Analysis (CDA), from which was created SCR was created a time signal sum of amplitude of phasic driver ($SCR_{AmpSum}(t)$) and area of the phasic driver ($SCR_{Area}(t)$), as shown respectively in equations 8 and 9. It used a sliding window ranging from 1 to 4 seconds after its time because it is a more conservative response window to capture the essential response after a stimulus [3].

$$SCR_{AmpSum}(t) = \sum_{k=1}^{4} peaks_{SCR}(t+k) \tag{8}$$

$$SCR_{Area}(t) = \sum_{k=1}^{4} Area_{SCR}(t+k) \tag{9}$$

2.5.7 Ocular measures

Eye tracker measurements were considered corrupted due to the incompatibility of the glasses and the crew's helmet, and thus, some segments of eye gaze recording were lost. Even though this problem, some portions of data could be analyzed in Tobii Lab Pro, which could be extracted from ocular measures: pupil dilatation and fixation heatmap.
Besides that, the video recorded from the eye tracker helped to achieve a better synchronization between flight and physiological data using maneuver segment events.

2.6 Discussion

2.6.1 Respiration activity

Respiration (RESP) frequency rate presented statistically significant a decreased during climb mission segment ($F(4,25) = 2.91, p = 0.042, \eta_g^2 = 0.32$), as present in Figure 6. The pairwise comparisons pointed out a reduction of respiration frequency rate during climb segment phase when compared with after shot moment (evade phase) (Table 2), a dropped off from range of 0.15-0.21 Hz to 0.13-0.16 Hz.
At the climb phase, the pilot is submitted by positive and negative G forces due to the rapid ascent from low-level flight to shot height and the attitude reversion towards the target that could explain this significant drop during this mission segment. Normal respiratory rate in humans is within the range of 10–20 breaths per min (0.16–0.33 Hz), and a range from 4 to 10 breaths per min (0.07–0.16 Hz) is considered slow breathing [15].

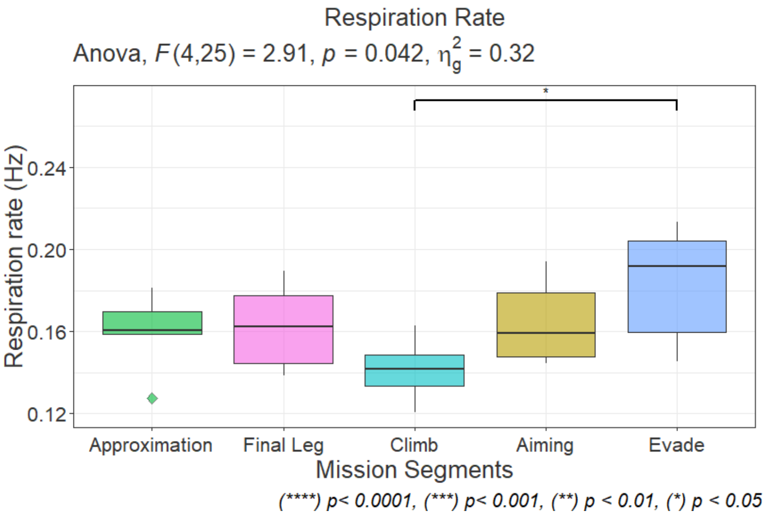


Figure 6 – Respiration rate per mission segments.

No difference was found between the other phases. Still, it could be verified in Figure 6 that respiration rate was maintained at around the low limit of the normal respiratory range. This type of breathing technique improves gas ventilation exchange and arterial oxygenation. It maintains parasympathetic–sympathetic balance and optimizes the cardiorespiratory reserve that could be called upon in times of intense physical or mental stress or activity [15]. The slow breathing pattern verified is on a marginal limit that could affect HRV measures [17].

Table 2 – Respiration frequency pairwise comparison per mission segment.

Comparisons	**Respiration frequency**		
	Mean [CI]	**p**	**d**
Final Leg - Approximation	-0.002 [-0.03;0.04]	0.87	-0.10(S)
Climb - Approximation	-0.02 [-0.05;0.02]	0.20	0.87(L)
Aiming - Approximation	-0.004 [-0.03;0.04]	0.87	-0.19(S)
Evade - Approximation	0.023 [-0.01;0.06]	0.20	-1.08(L)
Climb - Final Leg	-0.021 [-0.06;0.02]	0.20	-0.97(L)
Aiming - Final Leg	0.002 [-0.03;0.04]	0.87	-0.09(S)
Evade - Final Leg	0.021 [-0.01;0.06]	0.20	-0.98(L)
Aiming - Climb	0.023 [-0.01;0.06]	0.20	-1.07(L)
Evade - Climb	0.042 [0;0.08]	0.023 (*)	-1.96(L)
Evade - Aiming	0.02 [-0.01;0.06]	0.20	-0.88(L)

1: Turkey with Benjamini - Hochberg (BH) p adjustment.

2: [CI]: 95% Confidence Interval.

3: p-value conventions: (****) p< 0.0001, (***) p< 0.001, (**) p < 0.01, (*) p < 0.05.

4: Cohen's d conventions: Small(S) d< 0.5, Medium(M) 0.5< d < 0.8, Larger(L) d > 0.8.

On helicopter rocket firing was not verified the breathing rifle pattern [9], of shot systematically in a respiration break. This is also indicative that the pilots were in a regular breathing pattern. It could be understood that helicopter shot performance is less dependent on abdomen movement than rifleman shot, as aircraft has stability augmented systems that compensate for helicopter instability.

2.6.2 Electro cadiac activity

The analysis of heartbeat rate (HR) indicated statistically significant difference between mission segments ($F(4,25) = 62.69, p < 0.0001, \eta_g^2 = 0.91$), as shown in Figure 7. The heartbeat increases until the shot and after that starts to decrease; thus, physiological measures capture mental workload increase, as a state of body preparedness associated with increased activity in the nervous system

[14]. Pairwise comparisons per mission segment show large effect sizes (Table 3), revealing a pattern of rocket shot mission mental demand.

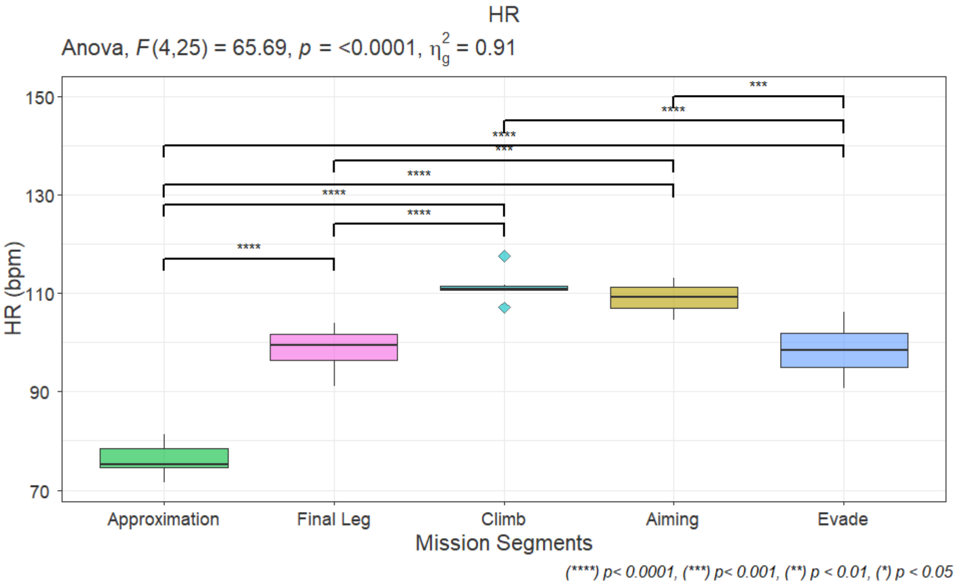


Figure 7 – Heartbeat rate per mission segments.

Table 3 – Heartbeat rate pairwise comparison per mission segment.

Comparisons	**Heartbeat**		
	Mean [CI]	**p**	**d**
Final Leg - Approximation	22.39 [15.25;29.53]	4.7 x 10^{-9} (****)	-5.32(L)
Climb - Approximation	35.24 [28.11;42.38]	1.1 x 10^{-12} (****)	-8.37(L)
Aiming - Approximation	32.86 [25.72;40.00]	2.7 x 10^{-12} (****)	-7.81(L)
Evade - Approximation	22.23 [15.10;29.37]	4.7 x 10^{-9} (****)	-5.28(L)
Climb - Final Leg	12.85 [5.71;19.98]	3.0 x 10^{-5} (****)	-3.05(L)
Aiming - Final Leg	10.47 [3.33;17.61]	2.8 x 10^{-4} (****)	-2.50(L)
Evade - Final Leg	-0.16 [-7.29;6.97]	0.94	0.04(S)
Aiming - Climb	-2.38 [-9.52;4.75]	0.37	0.56(M)
Evade - Climb	-13.01 [-20.15;-5.87]	3.0 x 10^{-5} (****)	3.09(L)
Evade - Aiming	-10.62 [-17.77;-3.49]	2.7 x 10^{-4} (****)	2.52(L)

1: Turkey with Benjamini - Hochberg (BH) p adjustment.

2: [CI]: 95% Confidence Interval.

3: p-value conventions: (****) p< 0.0001, (***) p< 0.001, (**) p < 0.01, (*) p < 0.05.

4: Cohen's d conventions: Small(S) d< 0.5, Medium(M) 0.5< d < 0.8, Larger(L) d > 0.8.

Regarding HRV time domain features, RMSSD and SD1 presented statistically significant difference between mission segments $(F(4,25) = 4.92, p < 0.005, \eta_g^2 = 0.44$ and $F(4,25) = 4.86, p < 0.005, \eta_g^2 = 0.44$ respectively), as shown in Figure 8.

The RMSSD was correlated $(r(30) = 0.88, p < 0.00001)$ with the non-linear metric SD1, as expected [17], and also either were able to indicate increased arousal at shot event surroundings.

This increase in mental workload was presented as a decrease in the beat-to-beat variance in HR, which reflects low HRV. RMSSD is considered a measure mediated by the vagal tone, which generally decreases during stress conditions [10], pairwise comparisons, presented in Table 4, showed a significant difference from the approximation phase from climb to shot mission segments.

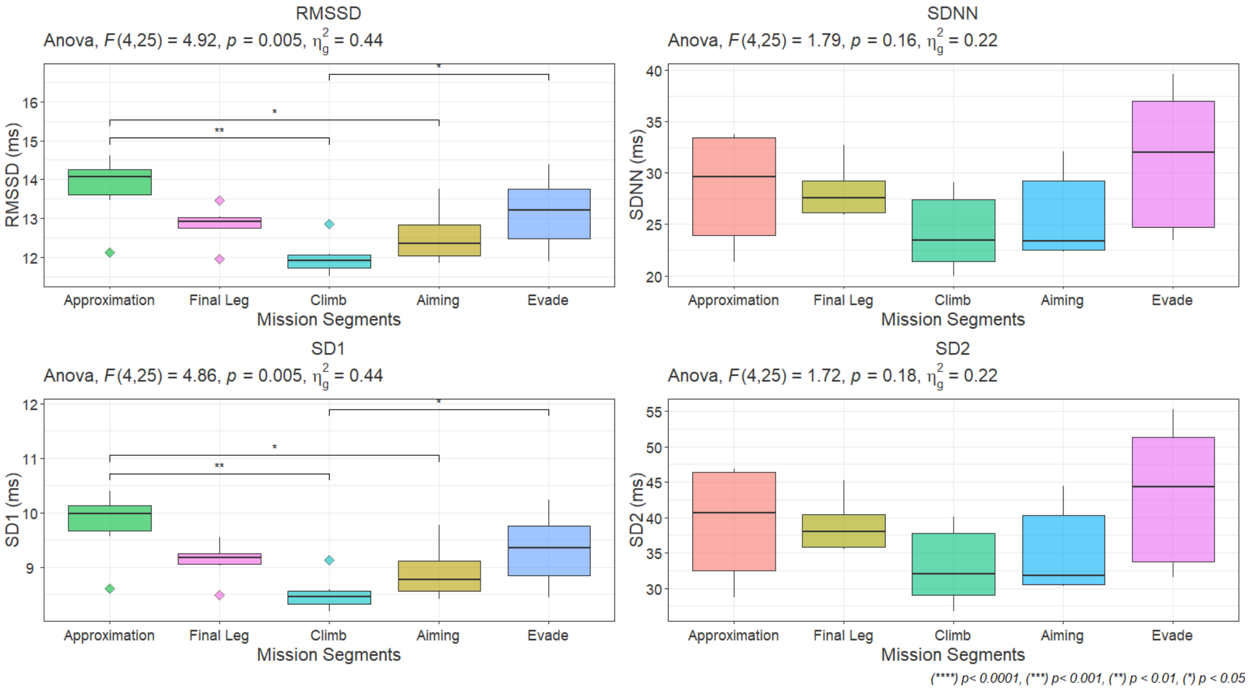


Figure 8 – HRV time domain parameters per mission segments.

Table 4 – HRV time-domain and Poincaré pairwise comparison per mission segment.

Comparisons	RMSSD			SDNN			SD1			SD2		
	Mean [CI]	p	d	Mean [CI]	p	d	Mean [CI]	p	d	Mean [CI]	p	d
Final - Approximation	0.93 [-2.18;0.31]	0.09	1.27 (L)	-0.29 [-8.80;8.21]	0.92	0.05 (S)	-0.67 [-1.56;0.21]	0.09	1.28 (L)	-0.29 [-12.65;12.07]	0.95	0.04 (S)
Climb - Approximation	-1.78 [-3.03;-0.53]	0.003 (**)	2.42 (L)	-4.28 [-12.79;4.22]	0.45	0.85 (L)	-1.26 [-2.15;-0.38]	0.003 (**)	2.41 (L)	-6.04 [-18.40;6.31]	0.46	0.83 (L)
Aiming - Approximation	-1.23 [-2.48;0.01]	0.038 (*)	1.68 (L)	-2.74 [–11.25;5.76]	0.50	0.54 (M)	-0.87 [-1.76;0.01]	0.038 (*)	1.67 (L)	-3.89 [16.25;8.46]	0.50	0.53 (M)
Evade - Approximation	-0.63 [-1.87;0.62]	0.23	0.85 (L)	2.84 [-5.66;11.35]	0.50	-0.56 (M)	-0.46 [-1.34;0.43]	0.24	0.87 (L)	4.07 [-8.28;16.43]	0.50	-0.55 (M)
Climb - Final Leg	-0.84 [-2.09;0.40]	0.12	1.14 (L)	-3.98 [-12.49;4.51]	0.45	0.79 (M)	-0.59 [-1.48;0.29]	0.12	1.13 (L)	-5.74 [-18.11;6.61]	0.46	0.79 (M)
Aiming - Final Leg	-0.30 [-1.54;0.94]	0.49	0.40 (S)	2.45 [-10.96;6.05]	0.51	0.48 (S)	-0.20 [-1.09;0.68]	0.50	0.39 (S)	-3.60 [-15.96;8.75]	0.50	0.49 (S)
Evade - Final Leg	0.31 [-0.94;1.55]	0.49	-0.42 (S)	3.13 [-5.37;11.64]	0.50	-0.62 (M)	0.21 [-0.67;1.10]	0.50	-0.40 (S)	4.36 [-8.00;16.72]	0.50	-0.60 (M)
Aiming - Climb	0.54 [-0.70;1.79]	0.26	-0.74 (M)	1.53 [-6.97;10.04]	0.67	-0.30 (S)	-0.39 [-0.50;1.27]	0.26	-0.74 (M)	2.14 [-10.21;14.51]	0.68	-0.29 (S)
Evade - Climb	1.15 [-0.09;2.40]	0.039 (*)	-1.57 (L)	7.12 [-1.38;15.63]	0.21	-1.42 (L)	0.81 [-0.08;1.69]	0.04 (*)	-1.54 (L)	10.11 [-2.24;22.47]	0.24	1.39 (L)
Evade - Aiming	0.61 [-0.64;1.85]	0.23	-0.82 L)	5.59 [-2.91;14.09]	0.33	-1.11 (L)	-0.42 [-0.47;1.31]	0.25	-0.80 (L)	7.97 [-4.39;20.32]	0.35	-1.09 (L)

1: Turkey with Benjamini - Hochberg (BH) p adjustment.

2: [CI]: 95% Confidence Interval.

3: p-value conventions: (****) p< 0.0001, (***) p< 0.001, (**) p < 0.01, (*) p < 0.05.

4: Cohen's d conventions: Small(S) d< 0.5, Medium(M) 0.5< d < 0.8, Larger(L) d > 0.8.

As shown in Table 4, SDNN and SD2 were not able to indicate any difference between mission segments, although other studies consistently indicate an SDNN decrease during stress conditions [10]. SDNN is strongly correlated with the frequency domain and is more influenced by changes in respiration [17]; in this case, the observed decrease in breathing rate could affect SDNN measures. Besides that, SDNN did not achieve acceptable correlations with short HRV when using ultra-short HRV [17]; by another hand, RMSSD is considered less affected by respiration and got better correlations results using ultra-short HRV.

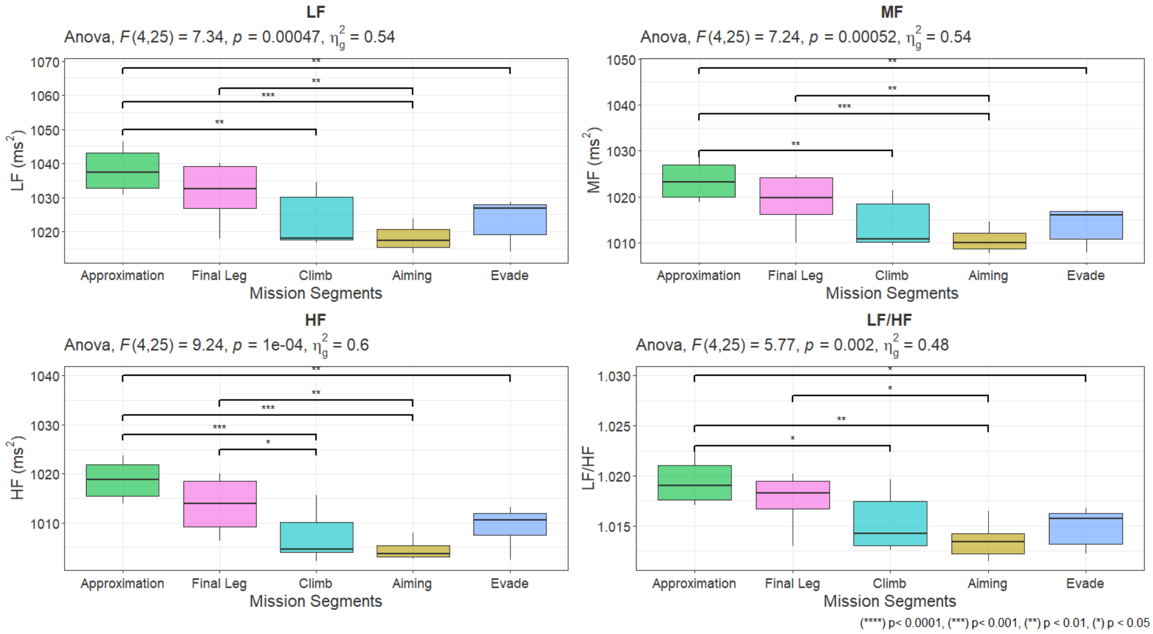


Figure 9 – HRV frequency domain per mission segments.

From the frequency domain data, as presented in Figure 9,LF, MF, HF bands, and LF/HF shifted to a lower level after the final leg phase, which is an indication of less parasympathetic activity during the climb to shot mission segment when compared with baseline approximation segment. From the literature, slow breathing rates could affect frequency domain features [17]. Still, it is unknown how respiratory activity could affect these measures, as breathing rate was in the marginal limit of normal respiration rate of around 0.15 Hz affecting all bands. Even though this fact, as presented in Table 5, climb and aiming mission segments were indicated with a significant increase in mental workload demands with a large effect size in coherence with time-domain HRV measures.

None of the HRV measurements could differentiate the mental workload between climb and aiming mission segments.

Table 5 – HRV Frequency domain pairwise comparison per mission segment.

Comparisons	**LF**			**MF**			**HF**			**LF/HF**		
	Mean [CI]	**p**	**d**	**Mean [CI]**	**p**	**d**	**Mean [CI]**	**p**	**d**	**Mean [CI]**	**p**	**d**
Final Leg-Approximation	-6.52 [-18.59;5.55]	0.18	0.91 (L)	-4.43 [-12.36;3.49]	0.16	0.94 (L)	-5.05 [-12.73;2.63]	0.11	1.11 (L)	-0.002 [-0.01;2.29]	0.26	0.74 (M)
Climb-Approximation	-14.88 [-26.95;-2.81]	0.005 (**)	2.09 (L)	-9.63 [-17.56;-1.70]	0.005 (**)	2.05 (L)	-11.60 [-19.28;-3.91]	0.0008 (***)	2.55 (L)	-0.004 [-0.01;-3.25]	0.016 (*)	1.71 (L)
Aiming-Approximation	-19.92 [-32.00;-7.85]	0.0005 (***)	2.80 (L)	-13.02 [-20.95;-5.09]	0.0005 (***)	2.78 (L)	-14.23 [-21.92;-6.55]	0.0001 (***)	3.14 (L)	-0.006 [-0.01;-1.83]	0.0025 (**)	2.46 (L)
Evade-Approximation	-14.51 [-26.58;-2.44]	0.005 (**)	2.04 (L)	-9.72 [-17.65;1.79]	0.005 (**)	2.07 (L)	-9.44 [-17.12;-1.75]	0.0043 (**)	2.08 (L)	-0.004 [-0.01;-4.62]	0.015 (*)	1.89 (L)
Climb-Final Leg	-8.36 [-20.43;3.71]	0.10	1.17 (L)	-5.19 [-13.12;2.73]	0.10	1.11 (L)	-6.54 [-14.23;1.13]	0.038 (*)	1.44 (L)	-0.002 [-0.01;1.73]	0.17	0.97 (L)
Aiming-Final Leg	-13.40 [-25.47;-1.33]	0.008 (**)	1.88 (L)	-8.58 [-16.51;-0.65]	0.009 (**)	1.83 (L)	-9.18 [-16.86;-1.49]	0.004 (**)	2.02 (L)	-0.004 [-0.01;-6.39]	0.016 (*)	1.72 (L)
Evade-Final Leg	-7.99 [-20.06;4.08]	0.10	1.12 (L)	-5.29 [-13.21;2.64]	0.10	1.13 (L)	-4.39 [-12.07;3.30]	0.13	0.96 (L)	-0.003 [-0.01;1.31]	0.12	1.15 (L)
Aiming-Climb	-5.04 [-17.11;7.04]	0.25	0.70 (M)	-3.39 [-11.31;4.54]	0.26	0.72 (M)	-2.63 [-10.31;5.05]	0.36	0.58 (M)	-0.002 [-0.01;2.26]	0.26	0.75 (M)
Evade-Climb	0.38 [-11.70;12.45]	0.93	-0.05 (S)	-0.09 [-8.01;7.84]	0.97	0.01 (S)	2.15 [-5.52;9.84]	0.41	-0.48 (S)	-0.0004 [-0.01;3.63]	0.75	0.18 (S)
Evade-Aiming	-5.41 [-6.66;17.48]	0.25	-0.76 (M)	-3.30 [-4.63;11.23]	0.25	-0.70 (M)	4.79 [-2.89;12.47]	0.11	-1.06 (L)	0.001 [-0.01;5.43]	0.36	-0.57 (M)

1: Turkey with Benjamini - Hochberg (BH) p adjustment.

2: [CI]: 95% Confidence Interval.

3: p-value conventions: (****) p< 0.0001, (***) p< 0.001, (**) p < 0.01, (*) p < 0.05.

4: Cohen’s d conventions: Small(S) d< 0.5, Medium(M) 0.5< d < 0.8, Larger(L) d > 0.8.

2.6.3 Electrodermal activity

The exam of the relationship between SCR features per mission segment and per pilot, as presented in Figure 10, where the shot event was compared with the mean of non-specific skin conductance response during each mission segment. It was identified a statistically significant difference $(F(5,30) = 2.9, p = 0.03, \eta_g^2 = 0.32)$ on SCR area per mission segment when compared with evade phase. Regarding SCR amplitude sum, it was verified a statistically significant difference $(F(5,30) = 4.15, p = 0.006, \eta_g^2 = 0.41)$ between aiming to shot event with others phases. As can be seen in 6, the SCR amplitude sum achieves a better phase differentiation than the SCR area, which could be related to the integration as a smooth operation was more affected by non-specific skin conductance response.

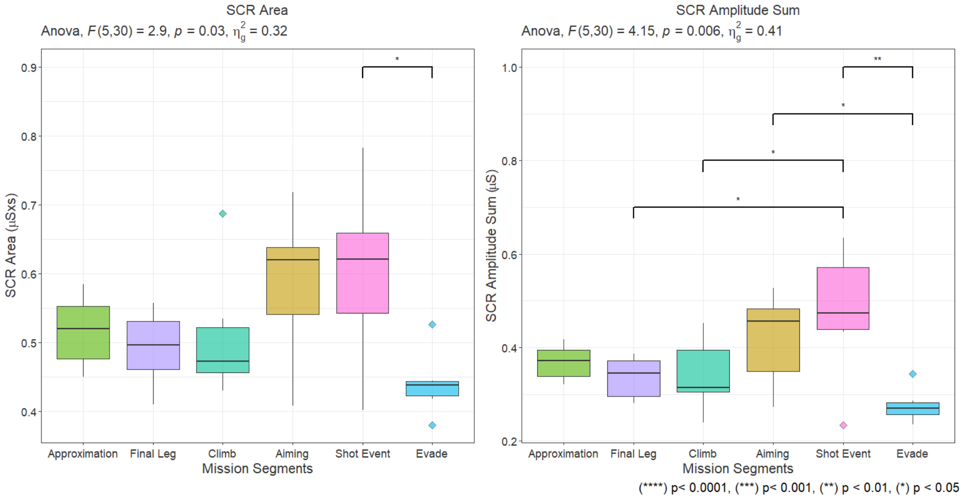


Figure 10 – SCR per mission segments.

These results of phasic component features have shown that EDA is sensitive and a valid indicator of arousal and mental workload, as cited in the literature [6]. At the aiming phase, SCR amplitude sum and area have shown significant correlation, (r(6) = 0.96, p < 0.0022), and both revealed a pattern of an increase of sympathetic activity in the surrounds of aiming to shot, marking this phase as the most stressful situations during the maneuver and identified a relief after the shot moment.

Table 6 – SCR pairwise comparison per mission segment.

Comparisons	**SCR area**			**SCR amplitude sum**		
	Mean [CI]	**p**	**d**	**Mean [CI]**	**p**	**d**
Final Leg - Approximation	-0.02 [-0.18;0.13]	0.78	-0.27(S)	-0.03 [-0.18;;0.11]	0.59	-0.36(S)
Climb - Approximation	-0.01 [-0.16;0.14]	0.87	-0.06(S)	-0.03 [-0.17;0.12]	0.60	-0.30(S)
Aiming - Approximation	0.07 [-0.08;0.22]	0.28	0.84(L)	0.05 [-0.01;0.2]	0.37	0.64(L)
Shot Event - Approximation	0.08 [-0.07;0.24]	0.25	1.03(L)	0.1 [-0.04;0.25]	0.11	1.36(L)
Evade - Approximation	-0.07 [-0.23;0.08]	0.27	-0.81(L)	-0.09 [-0.24;0.05]	0.16	-1.03(L)
Climb - Final Leg	0.02 [-0.14;0.17]	0.82	0.2(S)	0.004 [-0.14;0.15]	0.93	0.06(S)
Aiming - Final Leg	0.09 [-0.05;0.25]	0.22	1.11(L)	0.08 [-0.06;0.22]	0.20	1.0(L)
Shot Event - Final Leg	0.11 [-0.04;0.26]	0.18	1.3(L)	0.14 [-0.01;0.28]	0.035(*)	1.72(L)
Evade - Final Leg	-0.05 [-0.2;0.1]	0.44	-0.54(M)	-0.06 [-0.2;0.08]	0.33	-0.66(M)
Aiming - Climb	0.08 [-0.07;0.23]	0.27	0.9(M)	0.07 [-0.07;0.22]	0.21	0.94(L)
Shot Event- Climb	0.09 [-0.06;0.24]	0.21	1.1(L)	0.13 [-0.01;0.28]	0.035(*)	1.66(L)
Evade - Climb	-0.07 [-0.22;0.08]	0.28	-0.74(M)	-0.06 [-0.21;0.08]	0.32	-0.73(M)
Shot Event - Aiming	0.01 [-0.14;0.17]	0.83	0.19(S)	0.05 [-0.09;0.2]	0.36	0.72(M)
Evade - Aiming	-0.14 [-0.3;0.01]	0.05	-1.65(L)	-0.14 [-0.29;0.01]	0.035(*)	-1.67(L)
Evade - Shot Event	-0.16 [-0.31;-0.01]	0.049(*)	-1.84(L)	-0.19 [-0.34;-0.05]	0.004(**)	-2.39(L)

1: Turkey with Benjamini - Hochberg (BH) p adjustment.

2: [CI]: 95% Confidence Interval.

3: p-value conventions: (****) p< 0.0001, (***) p< 0.001, (**) p < 0.01, (*) p < 0.05.

4: Cohen's d conventions: Small(S) d< 0.5, Medium(M) 0.5< d < 0.8, Larger(L) d > 0.8.

2.7 Ocular measures

Due to the loss of eye gaze data, it was not possible to perform statistical analysis per mission segment of the ocular data. Even though this, it is possible to verify how the heatmap of a shot distributed the pilot's attention. In Figure 11 it is possible to confirm that there was a loss of calibration of the eye gaze point due to the interaction with the pilot´s helmet. As could be seen in Figure 11, before the aiming phase, the pilot's attention was divided between the location of the target and flight data parameters. The pilot's attention is entirely on the target in the aiming phase. This indicates increased mental workload because the pilot is wholly focused on a task.

(a) Prior aiming phase

(b) Aiming phase

Figure 11 – Eye tracker heatmap.

Observing the variation of the pupil diameter in Figure 12 during the aiming phase, it is possible to verify the gradual increase of this parameter after the reversion of aircraft attitude towards the target, which is associated with stressful situations. This sympathetic increase during the aiming phase indicated more arousal demands than the other mission segments complementing ECG and EDA analysis.

Besides the limitation of measuring pupil diameter outside a laboratory, the shots denoted in Figure 12 were done with a difference of fewer than five minutes at the same heading, so the illumination effects on the pupil size were the same.

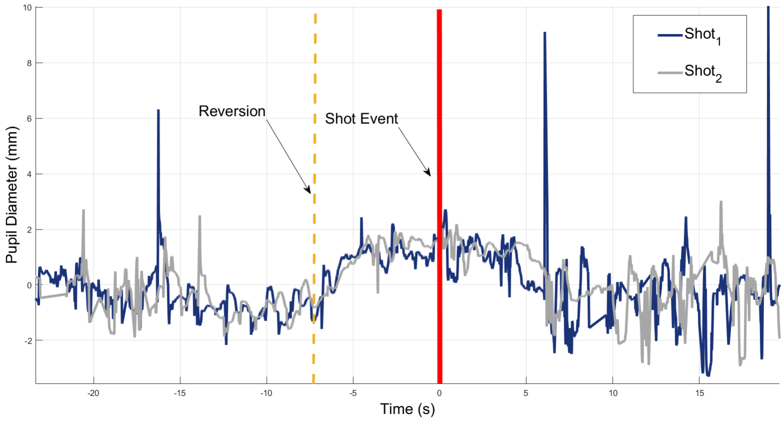


Figure 12 – Comparison of estimated impact point and real impact point on ground.

From the operational doctrine point of view, it is essential to emphasize that the second pilot should reinforce the management of flight parameters because the mission highly consumes the shooting pilot's attention demands and has a low response capacity to deal with unexpected events.

2.8 Performance versus physiological measures

The method proposed to estimate the rocket impact and the aiming path was validated by comparing the impact point estimated and the real impact on the target of the first rocket shot recorded by eye tracker camera, as shown in Figure 13, where it could be visualized that the impact was at the same area, using ground contour as reference. Even though this qualitative validation, the objective is to use the flight data as a metric to assess aircraft variations in the last two seconds before the shot event by computing aim holding box area [9], aiming path perimeter, and F-score [12].

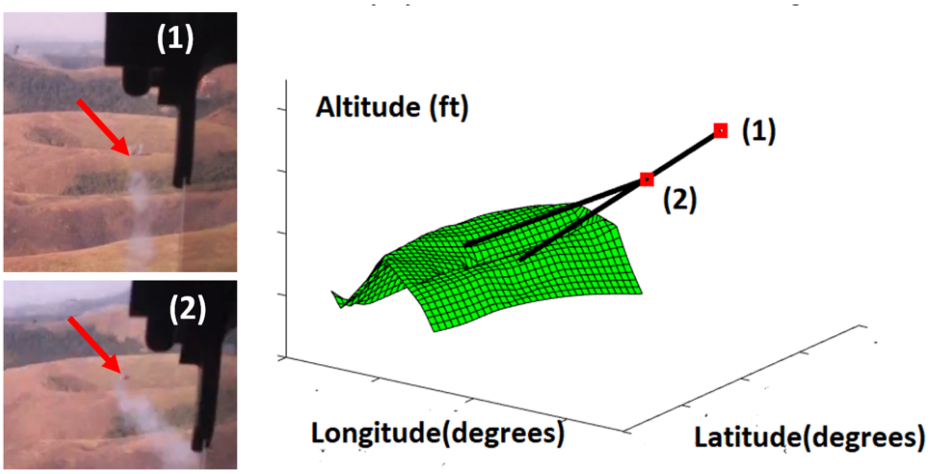


Figure 13 – Comparison of estimated impact point and real impact point on ground.

The relative relations among variations in performance and physiological features during the aiming phase were next analyzed with Pearson's correlation, according to Figure 14.

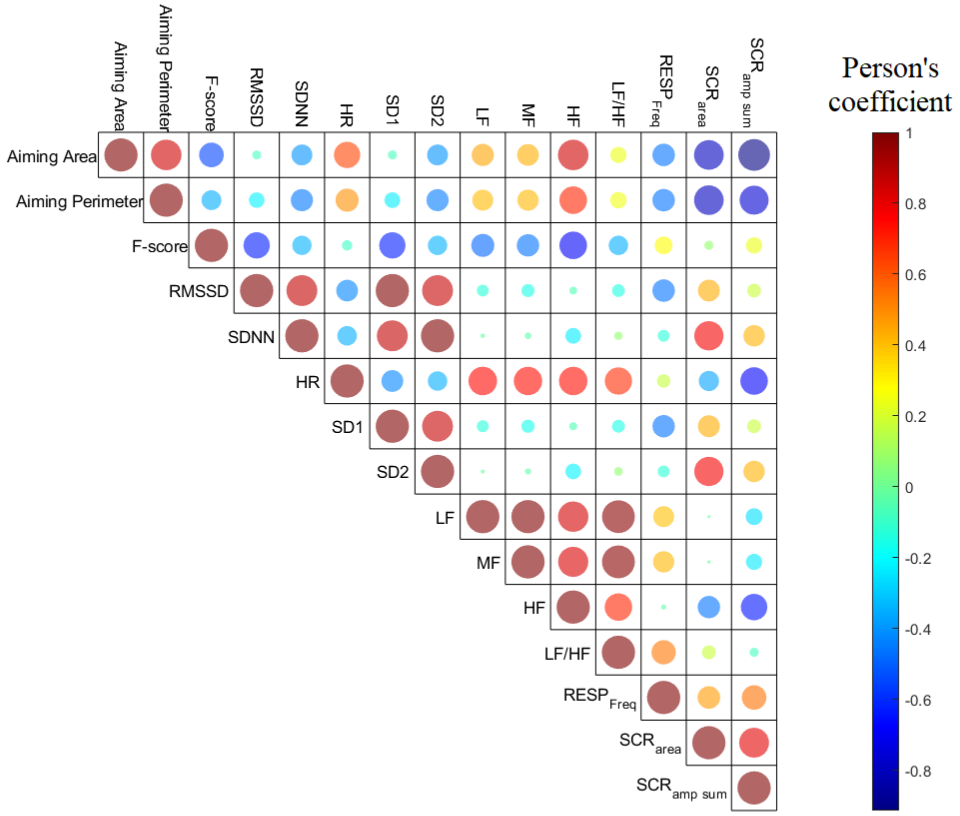


Figure 14 – Correlation matrix with performance and physiological at aiming phase.

Aim holding box area (aiming area) and aim path perimeter (aiming perimeter) presented a significant positive correlation ($r(6) = 0.85, p < 0.03$), meaning that when the pilot reduces the area actuating in-flight commands, it also minimizes the perimeter before pulling the trigger. Despite that, only the aiming area presented significant correlations with physiological data, which showed a significant negative correlation between the variations measured by SCR amplitude sum ($r(6) = -0.91, p < 0.01$) and positive correlation with the HF band ($r(6) = 0.84, p < 0.03$), which means that the task of shot in a narrow area increases sympathetic activity. On the other hand, F-score did not present any significant correlation with any physiological measure in the aiming phase, corroborating the idea that the pilot is more dedicated to keeping the aim towards the target and the importance of maneuver deviations was less important. From Figure 14 was visualized a strong and significant correlation between RMSSD-SD1 ($r(6) = 1, p < 0.0001$) and SDNN-SD2 ($r(6) = 1, p < 0.0001$), but none of them presented a significant correlation with the performance parameter. From mission segment analysis, the HF and SCR amplitude sum were also the parameters that better differentiated the aiming to shot period from the other mission phases.

Using principal component analysis (PCA) to reduce dimensionality is found that two eigenvectors explain 71.62% of the data variability and three eigenvectors 93.6% of data variability, thus PCA with three principal components will be addressed.

Eigenvectors are displayed in three dimensions in Figure 15, where it can be seen that the aiming area eigenvectors and aiming perimeter eigenvectors are near alignment with several ECG, SCR, and RESP vectors, and F-score is almost orthogonal to physiological eigenvectors. In this sense, this fact can concur with the same finding of ocular data that the pilot at the aiming phase did not pay attention to the deviations of the parameters maneuvers but was devoted to correcting horizontal attitude towards the target.

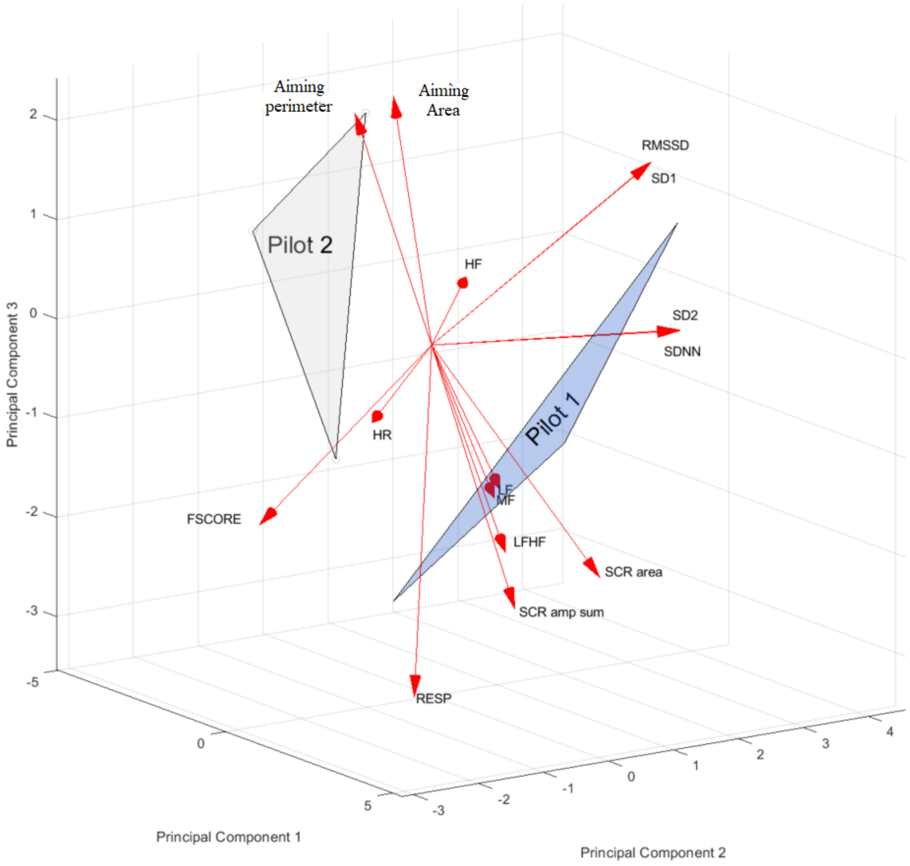


Figure 15 – Three dimensions principal component analysis during aiming phase.

A better understanding of the relations between these vectors was computed cosine similarity, as could be seen in Table 7. Results showed that the variations in aiming performance (area and perimeter) were more related to SCR amplitude sum, SCR area, and HF band. F-score got low similarity values with ECG parameters (RMSSD, SD1, and HF).

EDA data showed a more sensitive and consistent parameter to explain changes in mental workload from increasing task demand (aiming area and aiming perimeter). EDA features have shown a negative correlation with aiming area and perimeter, meaning that a more stable shot handling (lower area and perimeter) gets higher sympathetic activation measured by EDA features.

Both Pearson's correlation and PCA and HF band analyses appeared to barely correlate with performance metrics.

Analyzing the PCA clustered by pilots, in Figure 15, HF eigenvector direction is parallel relative to the pilot's planes, and SCR amplitude sum and area eigenvector directions are pointing towards Pilot 1. This presented a tendency that Pilot 1 was more mentally involved during the aiming than Pilot 2, by PCA analysis, which one of the influenced facts could be related to the fact has the last training of Pilot 1 was ten years before and Pilot 2 was one day. High mental demands could lead the pilot to commit errors in the proper target engagement or in the handling of any emergency event.

Table 7 – Eigenvectors cosine similarity of PCA.

	Aiming box area	Aiming path perimeter	F-score	HR	RMSSD	SDNN	SD1	SD2	LF	MF	HF	LF/HF	RESP rate	SCR area	SCR amp sum
Aiming box area		**0.99 [1]**	-0.61	0.51	0.15	-0.24	0.15	-0.24	0.22	0.20	**0.77 [4]**	0.01	-0.63	-0.78	**-0.97 [2]**
Aiming path perimeter	**0.99 [1]**		-0.47	0.53	-0.00	-0.40	-0.00	-0.40	0.18	0.16	**0.72 [5]**	-0.03	-0.59	**-0.86 [2]**	**-0.99 [1]**
F-score	-0.61	-0.47		-0.23	-0.78	-0.61	**-0.77 [3]**	-0.61	-0.44	-0.41	**-0.71 [6]**	-0.30	0.47	0.02	0.43
HR	-0.51	0.53	-0.23		-0.42	-0.37	-0.43	-0.37	**0.84 [3]**	**0.85 [3]**	**0.84 [1]**	**0.76 [3]**	0.35	-0.24	-0.40
RMSSD	0.15	0.00	**-0.78 [1]**	-0.42		**0.87 [2**]	1.00	**0.87 [2]**	-0.07	-0.10	0.11	-0.12	-0.55	0.27	-0.04
SDNN	-0.24	-0.40	-0.61	-0.37	**0.87 [2]**		**0.87 [2]**	**1.00 [1]**	0.15	0.13	0.01	0.19	-0.08	**0.70 [3]**	0.40
SD1	-0.15	0.00	**-0.77 [2]**	-0.43	**.00 [1]**	**0.87 [3]**		**0.87 [2]**	-0.07	-0.10	0.11	-0.13	-0.55	0.27	-0.03
SD2	-0.24	-0.40	-0.61	-0.37	**0.87 [2]**	**1 [1]**	**0.97 [1]**		0.16	0.13	0.01	0.19	-0.08	**0.70 [3]**	0.40
LF	0.22	0.18	-0.44	**0.84 [2]**	-0.07	0.15	-0.07	0.16		**1.00 [1]**	**0.79 [2]**	**0.98 [1]**	0.51	0.27	-0.02
MF	0.20	0.16	-0.41	**0.85 [1]**	-0.10	0.13	-0.10	0.13	**1.00 [1]**		**0.78 [3]**	**0.98 [1]**	0.54	0.27	-0.01
HF	**0.77 [3]**	**0.72 [3]**	**-0.71 [3]**	**0.84 [2]**	0.11	0.01	0.11	0.01	**0.79 [4]**	**0.78 [4]**		0.64	-0.08	-0.29	-0.61
LF/HF	0.01	-0.03	-0.30	**0.76 [3]**	-0.12	0.19	-0.13	0.19	**0.98 [2]**	**0.98 [2]**	0.64		0.67	0.43	0.19
RESP rate	-0.63	-0.59	0.47	-0.35	-0.55	-0.08	-0.55	-0.08	0.51	0.54	-0.08	0.67		0.62	**0.70 [4]**
SCR area	**-0.78 [4]**	**–0.86 [2]**	0.02	-0.24	0.27	**-0.70 [3]**	0.27	**0.70 [4]**	0.27	0.27	-0.29	0.43	0.62		**0.90 [3]**
SCR amp sum	**-0.97 [2]**	**-0.99 [1]**	0.43	-0.40	-0.04	0.40	-0.03	0.40	-0.02	-0.01	-0.61	0.19	**0.70 [1]**	**0.90 [1]**	

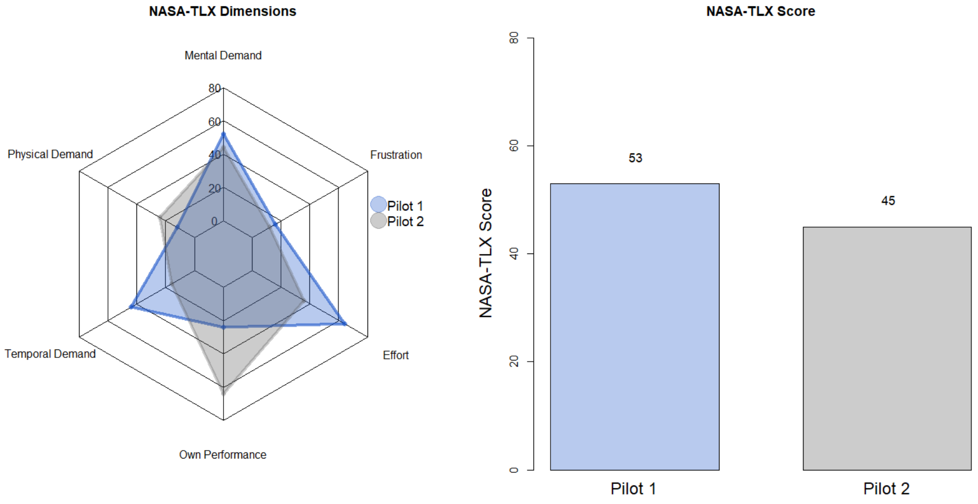
1: [ ] Cosine similarity ranking order in the column.

2.9 Subjective measures

Psycho-cognitive scales, which were assessed by questionnaires and NASA-TLX, were understood as a means of the workload and performance of all the maneuvers executed since the responses were collected after the flight. According to the questionnaires, the most stressful mission segment was the aiming phase due to the need to keep the aircraft's sighting system's aim steady on the target and meet the maneuver parameters.

The NASA-TLX score rated by the two pilots after the flight is presented in Figure 16. Task Load Index dimension analysis revealed that physical demand and frustration have a low contribution to overall workload rating for both pilots, as Figure 16(a) shows.

Pilot 1 score was higher than Pilot 2. This corroborates with the PCA output that Pilot 1 had more mental demand during the training. This also is reflected by higher effort and temporal demand ratings.



(a) (b)

Figure 16 – Comparison of the NASA-TLX between pilots.

3. Conclusion

The analysis of mental workload in real-life conditions is a complex task and involves the assessment of many human dimensions, which contributes to explaining different aspects. This article analyzed the operational helicopter maneuver of air-to-ground rocket-firing under the dimension of physiological measures through the NASA-TLX scale, ocular, electro cardiac, electrodermal, and respiration activities.

Mission segment observations have shown an increased HR pattern until the shot event and a shift down of HRV in time and frequency domains after the final leg approximation. Respiratory frequency rate was in the lower limit of normal respiration during mission segments, which is supposed to affect HRV features.

EDA feature of SCR amplitude sum has shown to be more sensitive to differentiate mission segments than SCR area.

Ocular measures have shown that the pilot's attention is devoted to pointing the rocket aim to the target during the aiming phase. An increase in pupil dilatation was observed during the aiming phase, reinforcing the idea of the aiming phase with greater mental workload demand.

The overall view is that the aiming phase is the more mental demand, indicated by ECG, EDA features, and total attention of eye fixation towards the target. HR was more sensitive to mission demands of flight because it shows more statistically significant effects among the flight segments.

Correlation analysis and principal component analysis have demonstrated that SCR area, SCR amplitude sum, and HF were the human dimensions more correlated with helicopter stability aiming performance before the shot event. More consistent indicated that phasic data EDA is more related to better performance. Conversely, the deviation of performance parameters (F-score) did not present any significant correlation with any physiological measure in the aiming phase, corroborating the idea that the pilot at this phase is more dedicated to keeping the aim towards the target, giving less importance to maneuver deviations.

Another important issue was that the NASA-TLX rating corroboration with the quantitative data, indicating the pilot with more mental workload demands and pilots' opinions also indicated the aiming phase as the most critical mission segment.

The results indicated that a less trained pilot would be more affected by mental demands and stress, thus being more successful in errors during the aiming phase. Indeed, pilots must do continuous practice sections to acquire the cognitive skill levels required. Still, as rockets may not be available all over the year for training, thus simulator would be an essential tool to reduce the pilot's mental demand for a real flight if a similar physiological pattern is observed in this environment.

Future work will be directed to measure the mental workload of a rocket shot in the simulator for comparison with real air-to-ground helicopter shooting.

4. Ethics statements

The experimental protocol was reviewed by the ethics committee with the global objective of the investigation of physiological data variables on pilot activity.

5. Copyright Statement

References

[1] Becerra-Luna B, Rosado R, Garcia J, et al. Heart Rate Variability Assessment Using Time–Frequency Analysis in Hypotensive and Non-Hypotensive Patients in Hemodialysis. *Applied Sciences* Vol.10(17), 6074, 2020. DOI: 10.3390/app10176074.

[2] Beggiato M, Hartwich F and Krems J. Using Smartbands, Pupillometry and Body Motion to Detect Discomfort in Automated Driving. *Frontiers in Human Neuroscience* Vol. 121, 2010. DOI: 10.3389/fnhum.2018.00338.

[3] Benedek M and Kaernbach C. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods* Vol. 190 (1), No. 30, pp. 80-91, 2010. DOI: 10.1016/j.jneumeth.2010.04.028.

[4] Boucsein W. Electrodermal Activity. *Springer* 2nd Ed, pp. 1071-1087, 2012. DOI: 10.1007/978-1-4614-1126-0.

[5] Brown S and Mitchell K. Shooting Stability: A Critical Component of Marksmanship Performance as Measured through Aim Path and Trigger Control. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* Vol. 61(1), pp. 1476-1480, 2017. DOI: 10.1177/1541931213601854.

[6] Charles R and Nixon J. Measuring mental workload using physiological measures: A systematic review. *Applied Ergonomics* Vol. 74, pp. 221-232, 2002. DOI: 10.1016/j.apergo.2018.08.028.

[7] Clifford G. Signal Processing Methods for Heart Rate Variability. *University of Oxford* 2002.

[8] Dahlström N, Nählinder S, Wilson G and Svensson E. Recording of Psychophysiological Data during Aerobatic Training. *International Journal of Aviation Psychology* Vol. 21(2), pp. 105-122, 2009. DOI: 10.1080/10508414.2011.556443.

[9] Espinosa P, Nagashima S, Chung, G, et al. Development of sensor-based measures of rifle marksmanship skill and performance. *University of California*, National Center for Research on Evaluation, Standards, and Student Testing. Report 756, pp. 21, 2009.

[10] Giannakakis G, Giannakaki K, Simantiraki O, et al. Review on psychological stress detection using biosignals, *IEEE Transactions on Affective Computing* Vol. 13, Issue 1, pp. 440-460, 2019. DOI: 10.1109/TAFFC.2019.2927337.

[11] Gospodinov M, Gospodinova E, Georgieva-Tsaneva G. Mathematical methods of ECG data analysis, *Academic Press*, Chapter 7, pp. 177-209, 2019. DOI: 10.1016/B978-0-12-815368-0.00007-5.

[12] Jin H, Hu Z, Li K et al. Study on How Expert and Novice Pilots Can Distribute Their Visual Attention to Improve Flight Performance. *IEEE* 2021. DOI: 10.1109/ACCESS.2021.3066880.

[13] Magnusson S. Similarities and Differences in Psychophysiological Reactions Between Simulated and Real Air-to-Ground Missions. *International Journal of Aviation Psychology* Vol. 12 (1), pp. 49-61, 2002. DOI: 10.1207/s15327108ijap1201_5.

[14] Roscoe A. Assessing pilot workload. Why measure heart rate, HRV and respiration?. *Biological Psychology* Vol. 34, pp. 259-287, 1992.

[15] Russo M, Santarelli D, Rourke D. The physiological effects of slow breathing in the healthy human. *Breathe* 13, pp. 298-309, 2017. DOI: 10.1183/20734735.009817.

[16] Scarpari J, Ribeiro M, Deolindo C et al. Quantitative assessment of pilot-endured workloads during helicopter flying emergencies: an analysis of physiological parameters during an autorotation. *Sci Rep* 11, 17734, 2021. DOI: 10.1038/s41598-021-96773-y.

[17] Shaffer F, Ginsberg J. An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health), 5* Vol. 5, 2017. DOI: 10.3389/fpubh.2017.00258.

[18] Shaffer F, Meehan Z, and Zerr C. A Critical Review of Ultra-Short-Term Heart Rate Variability Norms Research. *Frontiers in Neuroscience* Vol. 14, 2020. DOI: 10.3389/fnins.2020.594880 .

[19] Vest A, Da Poian G, Li Q et al. An open source benchmarked toolbox for cardiovascular waveform and interval analysis. *Physiological Measurement* Vol. 39(10), 2018. DOI: 10.1088/1361-6579/aae021.

[20] Wilson G. Air-to-ground training missions: A psychophysiological workload analysis. *Ergonomics* Vol. 36(9), pp. 1071-1087, 1993. DOI: 10.1080/00140139308967979.