

Received December 13, 2021, accepted January 18, 2022, date of publication January 20, 2022, date of current version February 14, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3145043

# Human Heart-Related Indexes Behavior Study for Aircraft Pilots Allowable Workload Level Assessment

ANDREA ALAIMO<sup>1</sup>, ANTONIO ESPOSITO<sup>1</sup>, PALMIRA FARACI<sup>2</sup>, CALOGERO ORLANDO<sup>1</sup>, AND GIUSY DANILA VALENTI<sup>2</sup>

<sup>1</sup>Faculty of Engineering and Architecture, Kore University of Enna, 94100 Enna, Italy

<sup>2</sup>Faculty of Human and Social Sciences, Kore University of Enna, 94100 Enna, Italy

Corresponding author: Antonio Esposito (antonio.esposito@unikore.it)

This work was supported by the Italian Ministry of Education, University and Research Department for Higher Education and Research Directorate General for the Coordination, Promotion and Valorisation of Research Notice for the presentation of industrial research and experimental development projects in the 12 Specialization Areas identified by the 2015–2020 NRP. Project “RPASInAIR - Remotely Piloted Aircraft Systems In unsegregated AIRspace for services” - ARS01\_00820-CUP J66C18000460005.

This work involved human subjects. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Committee of the Kore University of Enna.

**ABSTRACT** This study aimed to evaluate workload by detecting Heart Rate Variability (HRV) indexes in a sample of 34 pilots (with a mean age of 33 years) while performing simulated flight exercises. A one-way ANOVA with repeated measures was performed to assess the changes of the physiological measures in five standard maneuvers associated with different workload levels. The results show that all the indexes, but the Low Frequency to High Frequency ratio index (LF/HF), have a well-defined trend between the baseline and the en-route phase and with the three phases takeoff, steady turn, and landing. This study, as main findings, provides evidence of a differentiation among low, medium, and high workload levels using the time, frequencies, and non-linear HRV domains of analysis. These findings support the relevance of HRV indexes for workload evaluation, suggesting the development of non-invasive instruments capable of assessing workload in real-time. Further studies may be conducted to investigate whether the same findings could also be applied to more challenging maneuvers in real working conditions.

**INDEX TERMS** Pilot performance monitoring, heart rate variability, workload assessment.

## I. INTRODUCTION

The handling qualities of the aircraft are related to the maneuvers workload level experienced by the pilot while performing a specific task during the different flight phases [1]. A well-known consistent method to quantitatively define handling qualities is the Cooper-Harper pilot rating scale [2]. Using a decision-making process, the pilot can assign a rate of the aircraft's controllability and accuracy in performing a given task. Regardless of the aircraft attitude, the value of the rating scale achieved depends on the skill of the pilot and on subjective qualitative judgments or perceptions. Different ratings can be assigned to the same aircraft under the same flight conditions based on different pilot's abilities and backgrounds. However, recently, the idea of assessing both workload and the handling qualities using only classical

questionnaire-based methods has been revised by also taking into account measurements of a physiological nature with the aim of quantifying pilots' workload response. To do this, flight simulators for the analysis of the pilots' response to workload levels and tasks have been increasingly used in recent years [3], [4]. Motivation is found in the economic and safety advantages of training activities compared to a real flight, especially in the case of military pilots. For example, the studies carried out by Mansikka *et al.* [5] are focused on data collected from military pilot samples in training with particular reference during instrumental approaches. Lehmann *et al.* [6] analyzed the impact of turbulence and degraded visual environment on the pilot workload. Mohanavelu *et al.* [7] analyzed missions carried out on a simulator with different workload levels in terms of performance and additional tasks. Fuentes-García *et al.* [8] examined pilots performing a real flight with an F5 aircraft and a simulated flight with an operational F-5M flight simulator

The associate editor coordinating the review of this manuscript and approving it for publication was Eunil Park<sup>1</sup>.

by analyzing the variations in objective parameters related to heart rate variability. Klyde *et al.* [9] ran tests with a flight simulator and used electroencephalogram (EEG) and electrocardiogram (ECG), as physiologically based techniques, to estimate handling qualities. Their results also demonstrated the feasibility of artificial neural networks in using different physiological measurements.

Another field where the pilot mental workload analysis and modeling can be beneficial regards the pilot response models related to the so-called pilot-in-the-loop investigation. Thus, the studies on pilot workload behaviors can improve the accuracy of the models used to analyze human assessment within complex systems, not only in the case of fixed wings [10] but also for missions with helicopters or for space missions. Another advantage is the possibility of adapting the levels of automation of remotely piloted aircraft systems during flight phases characterized by disturbances or systems faults. According to Ji *et al.* [11], it is crucial to investigate the pilot workload to improve flight safety and performance during different operations. The pilot model usually takes into account the neuromuscular behavior associated with the pilot actions along with the time delay to account for the pilot's cognitive responsiveness, while the pilot's mental workload due to physiological aspect could improve the pilot-model-in-the-loop [12].

However, it should be noted that an important issue to be tackled concerns the relationship between subjective and objective (physiologically based) measurements because a useful pilot model is not achievable without an adequate correlation on workload levels. Many authors have debated this issue in recent years. In a large number of studies [13], [14] workload level was evaluated using subjective measures such as the NASA-TLX and the Cooper-Harper scale. In some cases, it is possible to associate the pilots' perception of workload levels with an increase or decrease in the cardiac variability indexes. However, there is no consistency among the different authors. For example, comparing the Cooper-Harper scale with the NASA-TLX subjective test and HRV, Mansikka *et al.* [15] showed that HRV indexes fail to sufficiently capture variations between low and medium workload levels. Moreover, Alaimo *et al.* [16] have found that the pilot's subjective judgment cannot properly be associated with the human body's physiological results, since the pilot perception subjective may jeopardize the judgment in the performed operations.

In this framework, the objective measurements aim to assess the experienced workload level using only physiological parameters. Thus, they represent useful techniques for workload evaluation, and they are widely applied in several studies and researches [17]–[21].

Physiological measurements in the aviation field have some advantages. First, these measurements are widely tested and used in the medical sector, and their validity and reliability are not a debatable issues. Second, they offer an objective assessment, independent of individuals' own subjective perception, by providing information regarding their physical

state. In addition, the detection of physiological parameters can allow a workload assessment in real-time, leading to a continuous monitoring of changes in the amount of physical and mental effort experienced during the performed activities [22]. By doing so, it is easier to identify which specific task induces an increase or decrease in workload. An additional strength of these techniques lies in the use of wearable or contactless sensors with a low degree of intrusiveness with the performance [23], [24]. However, some drawbacks characterize this type of measurement: it is difficult to use standardized pilot groups to compare the obtained values; their variability among individuals could be a limitation for low-size samples; some of them are obtrusive, requiring special equipment and training, and there is not a single measure that can be universally valid in assessing workload across various scenarios because physiological responses caused by workload are highly scenario-dependent [25]–[27]. However, these limitations motivate further research in these directions.

In particular, the current study first examines the effects of demographic variables on HRV indexes. Then, by assuming that different maneuvers imply different workload levels for pilots, a relationship between flight phases and HRV indexes is examined by employing the ANOVA approach. Furthermore, we also assume that the HRV indexes show statistical differences among low, medium, and high levels of workload. To this purpose, a multivariate analysis is then performed.

## II. HRV MEASUREMENTS

In healthy and not stressful conditions, the intervals between consecutive heartbeats are not constant but vary randomly from one beat to another. Heart Rate Variability (HRV) is the natural variation in the time between one beat and the next one. It is also known as RR variability, where RR is the distance between two R peaks of ECG.

HRV analysis is a method used to assess the mechanisms of regulation of the human body's physiological functions. These mechanisms refer to the autonomic nervous system and the neuroendocrine system. The balance of these systems determines the capacity and type of adaptation to an external stimulus, which is commonly called a reaction. The adaptation, whether positive or negative, depends on the degree of irregularity of these mechanisms.

The HRV indexes are closely related to the autonomic nervous system, they can reflect the state of psychological stress in people, and are sensitive to changes related to the demands of the current activity and the effort required; therefore, they are widely used as indicators of workload [4], [8], [28]–[32].

It should be noted that although the study of HRV originates from a purely medical context and it is still analyzed and evaluated in clinical studies, its use in aviation is valid and demonstrated by several studies [10], [15], [16], [28], [30], [32]–[41].

HRV can be evaluated using different techniques. The most common are: a) the time domain analysis, b) the frequency domain (or power spectral density) analysis, and c) the non-linear indexes analysis. In the following paragraphs, the

measures extracted from the signals acquired on the sample under examination are described in detail.

### A. TIME-DOMAIN INDEXES

The time-domain analysis provides indexes that quantify the heart rate variability by using statistical approaches and geometric measurements (related to the density of distribution of RR values). Since these indexes are calculated with procedures that do not depend on particular cases of application, provided that the duration of the acquisitions is the same, it is possible to consider that the data are comparable with each other even if collected in research carried out in different times and contexts.

The main time-domain measurements include the computation of Mean RR, the standard deviation of the “normal to normal” series of intervals (SDNN), the square root of the mean squared differences of successive RR series (RMSSD), the number of successive differences that are greater than  $x$  milliseconds (NN $x$ ), and the percentage of total intervals that differ successively by more than  $x$  milliseconds (pNN $x$ ). For more details about the definition of such indexes the interested reader is referred to [42]. In this study, the RMSSD and the SDNN were considered. In particular, RMSSD is an estimate of the short-term components of HRV; SDNN, measured in milliseconds, reflects the influence of all nervous system-related factors contributing to HRV. It is worth noting that SDNN values depend on age and sex, and these factors should be considered whenever possible; of course it depends on the available population. A reduction in SDNN suggests activation of sympathetic regulation that inhibits the autonomic cycle. A strong reduction in SDNN results from a critical strain on regulatory systems and the involvement of higher levels of control and can be associated with an increase in workload level [42], [43].

An additional parameter considered in the present study is the one introduced by Baevsky in the space field for research on ECG signals of astronauts before and after space missions [44], [45]. The Baevsky stress index, hereafter referred to as SI, is an index that reflects the degree of centralization of heart rate control and essentially characterizes the activity of the sympathetic part of the autonomic nervous system. In this work, the square root of the Baevsky stress index was computed according to [46]. Of course, it should be noted that SI depends on the action accomplished. It makes a difference if you are sitting quietly or exercising. However, at least in general terms, the lower the SI value, the greater the number of difficult tasks a pilot can execute. The less variable the RR intervals are, the higher the SI scores will be.

Additional measurements of the geometrical characteristics are associated with the time domain RR variability. Indeed, starting from the RR histogram, it is possible to extract two measurements [42], the HRV triangular index ( $HRV_{ti}$ ) and the triangular interpolation of the NN interval histogram (TINN). The HRV triangular index measurement is the integral of the density distribution divided by the maximum value of the density distribution. TINN is the base

width of the distribution by approximating the NN interval distribution with a triangle.

### B. FREQUENCY-DOMAIN INDEXES

Power spectral analysis allows obtaining information regarding the frequency and amplitude of each specific rhythm present in the ECG. The power spectrum can be divided into four frequency bands [42]: i) HF (High Frequency) frequencies between 0.15 and 0.4 Hz; ii) LF (Low Frequency) frequencies between 0.04 and 0.15 Hz; iii) VLF (Very Low Frequency) frequencies between 0.003 and 0.04 Hz; iv) ULF (Ultra Low Frequency) the band falls below 0.003 Hz. The HF band is mainly considered an expression of the activity of the Parasympathetic Nervous System (PNS) (Sztajzel, 2004). The LF band is considered mainly due to the activity of the Sympathetic Nervous System (SNS) [47]. The VLF band is only partly due to SNS activity [48]. The oscillations in heart rhythms within the ULF band are mainly due to very slow regulatory processes, such as regulating body temperature, metabolism, and circadian oscillations. Thus, given the very low frequency of the oscillations, the ULF band's contribution can only be appreciated in 24-hour acquisitions [49].

Considering that observation time intervals equal to 5 minutes are used in the present work, and according to the bands' classification of the cardiology task force [42], it is decided to use only the three bands VLF, LF, and HF. Moreover, the total power ( $TOT_{pow}$ ), which is the sum of the energy in the VLF, LF, and HF bands for short-term recordings, is considered.

The LF/HF ratio, which is the power in the LF band divided by the power in the HF band [50] is considered an index of the global sympathovagal balance [47]. This ratio is very complex, since both sympathetic and parasympathetic activities are present in the LF band; therefore, it cannot be seen, in an absolute way, as a balance between the activities of the two categories of the autonomous system. This complicates the estimation of workload conditions; in fact, in some works, LF/HF increases as the workload increases [51]–[53]; whereas it decreases for Hsu *et al.* [54] it decreases. Considerations based on this relationship must be made knowing the LF band generation mechanism and the context in which the acquisitions are made [55].

### C. NON LINEAR INDEXES

Nonlinear indexes used in HRV analysis include Poincaré plot, detrended fluctuation analysis, approximate entropy, and sample entropy calculation [56]. In the present work, the parameters extracted from Poincaré's graphical method are used. In the Poincaré diagram, the scattering of RR $n$  (the time interval between two consecutive R peaks) over RR $n+1$  (the time between the two successive R peaks) is plotted [57]. An ellipse with a semi-major axis along the bisector of the axes plot can be reproduced on the graph. The standard deviations along the bisector identity line (SD2) and perpendicular to the bisector (SD1) represent the amplitudes of the major and minor axes of the ellipse, respectively. SD1 represents the standard deviation of the instantaneous

TABLE 1. HRV indexes.

Index	Description	Trend with increasing workload
SI	Root square of Baevsky stress index	Increases
SDNN	Standard Deviation of Normal-to-Norma RR intervals	Decreases
RMSSD	Root Mean Square of Successive Differences RR interval	Decreases
$HRV_{ti}$	HRV triangular index	Decreases
TINN	Triangular Interpolation of Normal-to-Normal intervals	Decreases
VLF	Very Low Frequency (frequency band within 0-0.04 Hz)	Decreases
LF	Low frequency (frequency band within 0.04-0.15 Hz)	Decreases
HF	High frequency (frequency band within 0.15-0.4 Hz)	Decreases
$TOT_{pow}$	Total power (sum of the VLF, LF, and HF bands)	Decreases
LF/HF	Ratio between LF and HF bands	Increases
SD1	Long term HRV of Poincare' plot	Decreases
SD2	Short term HRV of Poincare' plot	Decreases

beat-to-beat variability or short-term variability. SD2 represents the standard deviation of continuous or long-term variability [57]–[59]. The ellipse is primarily a visual aid and the numerical values of the SD1 and SD2 standard deviations contain the important data. Low values of SD1 indicate high levels of workload; in fact, this parameter is closely related to the SDNN baseline statistical measures.

Table 1 provides a summary list of the indexes considered in the subsequent analyses.

### III. METHOD

#### A. PARTICIPANTS AND EXPERIMENTAL SET-UP

In this study, different flight phases, classified as Category B and C according to the taxonomy described in the Military Specification 8785C [60] are compared by using ECG indexes based on HRV measurement that are acquired during exercises carried out by using a ground-based full flight simulator. The aim is to evaluate how the trend of the physiological indexes change in the different flight phases. An EcgMove 3 sensor manufactured by Movisens GmbH with a sampling frequency of 1024Hz was used for the ECG data acquisition. 34 participants (29 males and 5 females) aged between 23 and 52 years ( $M = 33$ ,  $SD = 8.37$ ) are considered for this experiment. Their BMI was in a healthy average range for both males and females ( $M_{males} = 25.18$ ,  $SD_{males} = 2.71$ ;  $M_{females} = 22.96$ ,  $SD_{females} = 1.96$ ). None of them declared to have any diseases, illness, or particular conditions that could alter the responses to physiological measures. It is worth noting that a total of 37 pilots participated in the experiments, however, data belonging to three of them were found to be outliers and thus were excluded from the analysis, resulting in a population of 34 pilots. Each pilot signed an informed consent form before the start of the exercise. The experimentation was carried out by means of a ground based full flight simulator that replicates the business jet Cessna Citation 560 XLS. Figure 1 shows an example of a flight path considered in the analysis. The flight phases performed by

each pilot under the same instructor directives are shown in blue, whereas the red dashed lines are connections between flight phases not evaluated in the present study. The authors chose the acronyms of different segments according to the ICAO ADREP Taxonomy. STD stands for the standing phase, the aircraft is on the runway head, and the pilot is staying for five minutes in a rest time seated position; TOF stands for Take-off; MNV is the flight phase associated with the turn maneuver, all the pilots performed two steady turns, the second one was in the opposite direction to the first; ENR was a segment of five minutes associated with level flight attitude; finally, the LDG is the landing flight phase. The segments length was selected based on a time interval of five minutes in order to compare the phases according to the minimum time recommended for HRV analyses [42]. It is worth noting that, concerning the handling quality classification [60], the terminal flight phases TOF and LDG are classified as Category C (since usually require accurate flight–path control), whereas ENR belongs to Category B nonterminal flight phases (which are normally accomplished using gradual maneuvers and without precision tracking). This gives an indication of the difficulty associated with the flight phases, and thus it can provide insight into the pilots' workload levels.

#### B. DATA ANALYSES

Before running the main statistical analyses, the first step is to make sure that our sample is in a resting state during rest time. To this purpose, a large set of physiological indexes during rest time was measured and compared with data obtained from other sources [61], [62], where these indexes were assessed in a quiet situation among healthy individuals. Data were obtained from open access Physionet database, an on-line collection of physiological and clinical data [63].

Mean scores on physiological indexes obtained from the total sample of 33 individuals were then considered as reference standards for the present study. To accomplish this goal, t-test independent samples is performed. To assess whether scores on the physiological indexes could be affected by demographic variables (sex and age), analyses of variances were conducted, taking into account mean scores obtained by the two subgroups (males vs. females,  $\leq 30$  years vs.  $> 30$  years, respectively).

To evaluate the differences among scores on physiological measures at the five flight phases, a ONE-way ANOVA with repeated measures was conducted. A mixed designed ANOVA with repeated measures was then performed to assess whether changes in phases were invariant across the two different age groups. The Greenhouse-Geisser correction is applied for the degrees-of-freedom adjustment. Descriptives, F values, and effect size ( $\eta^2$ ) were reported. A 5% significance level was adopted for all tests.

All the indexes were computed with the Kubios HRV – Heart rate variability analysis software [46], the data are processed using IBM SPSS Statistics (Version 20.0).

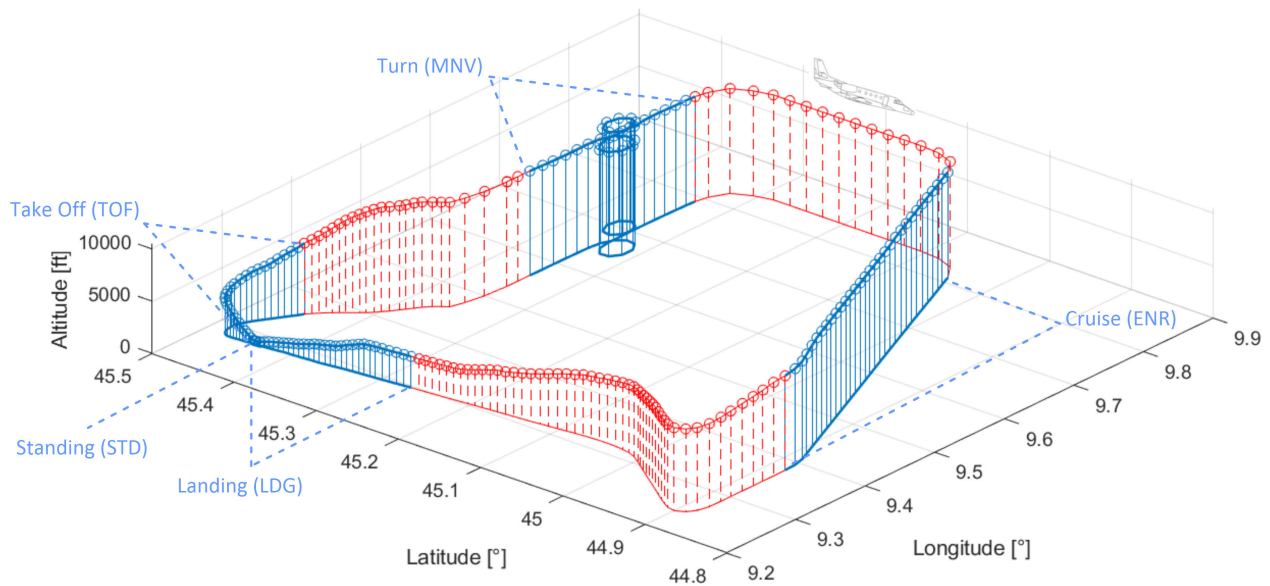


FIGURE 1. Example of flight phases taken into account for the analysis.

TABLE 2. Descriptives and the results from t-test.

Index	Group	M	SD	t	df	p
SI	1	12.773	4.581	0.426	57.703	0.672
	2	12.138	7.267			
SDNN	1	41.696	15.013	-0.275	61.497	0.784
	2	42.834	18.561			
RMSSD	1	30.542	17.641	-1.634	64.872	0.107
	2	37.637	17.892			
HRV <sub>ti</sub>	1	9.509	3.283	-1.299	60.675	0.199
	2	10.704	4.182			
TINN	1	224.029	78.273	0.298	62.108	0.766
	2	217.696	94.481			
VLF	1	86.613	70.054	1.309	64.987	0.195
	2	64.695	67.005			
LF	1	1206.147	807.103	-0.007	45.187	0.995
	2	1208.413	1716.676			
HF	1	520.407	588.791	-1.461	57.878	0.149
	2	776.334	822.218			
TOT <sub>pow</sub>	1	1813.685	1319.878	-0.574	55.596	0.569
	2	2049.903	1976.506			
LF/HF	1	4.163	2.775	0.458	47.661	0.649
	2	3.683	5.367			
SD1	1	21.628	12.495	-1.634	64.872	0.107
	2	26.652	12.673			
SD2	1	54.457	18.425	0.143	59.306	0.887
	2	53.698	24.586			

Note: Group 1 = Current sample; Group 2 = Physionet sample.

IV. RESULTS

A. REST TIME ANALYSIS

Table 2 shows descriptives and the results from t-test analyses for all the selected HRV indexes. Group 1 is the sample of pilots, whereas group 2 is the sample of Physionet data. Results indicate that the two samples do not differ each other. As a matter of fact, by applying the t-test, all the comparisons show a p value greater than .05, suggesting that no statistical differences are found between the two groups. This means that participants of the current sample are actually in a rest

condition during rest time. In addition, we can also assume that our sample is made up by healthy individuals.

B. RELATIONSHIPS BETWEEN SEX AND PHYSIOLOGICAL INDEXES

Table 3 lists the results of the ANOVA test for grouping based on gender. The results show that only scores on the Stress Index (SI) differed between males and females. This means that, except for this index, gender does not affect scores on physiological measurements, at least, considering the present sample.

C. RELATIONSHIPS BETWEEN AGE AND PHYSIOLOGICAL INDEXES

Regarding the associations between age and scores on physiological measurements, the sample was divided into two groups composed of individuals older than 30 years and equal or less than 30 years, respectively. The results in Table 4 show that only HF and LF/HF report significant differences between the two groups. The former is higher in younger individuals, whereas the latter increases with aging. Besides, LF/HF seems to be the most sensitive index to the age effect since it reports the highest effect size.

D. ONE-WAY ANOVA WITH REPEATED MEASURES ON DIFFERENT FLIGHT PHASES

In this section, the results of the one-way ANOVA with repeated measures on different flight phases are reported. The findings of the analyses are shown in Table 5. After checking the univariate tests, all but the LF/HF report a significant effect (p value lower than 0.001), suggesting that there is a

TABLE 3. ANOVA between sex.

Index	Sex	M	SD	F	p	$\eta^2$
SI	Males	12.092	4.159	<b>4.865</b>	0.035	0.132
	Females	16.721	5.406			
SDNN	Males	43.47	14.916	2.911	0.098	0.083
	Females	31.41	12.121			
RMSSD	Males	32.554	18.081	2.698	0.11	0.078
	Females	18.871	8.902			
HRV <sub>ti</sub>	Males	9.767	3.278	1.227	0.276	0.037
	Females	8.012	3.232			
TINN	Males	234.552	76.732	3.874	0.058	0.108
	Females	163	62.229			
VLF	Males	87.035	73.556	0.007	0.934	0
	Females	84.168	51.047			
LF	Males	1268.95	830.582	1.201	0.281	0.036
	Females	841.905	592.172			
HF	Males	595.291	612.752	0.345	0.561	0.011
	Females	376.084	448.649			
TOT <sub>pow</sub>	Males	1901.81	1371.83	0.876	0.356	0.027
	Females	1302.58	903.225			
LF/HF	Males	4.046	2.655	0.341	0.563	0.011
	Females	4.839	3.685			
SD1	Males	23.053	12.807	2.698	0.11	0.078
	Females	13.361	6.303			
SD2	Males	56.566	18.203	2.716	0.109	0.078
	Females	42.231	16.191			

Note: significant differences are in bold.

statistically significant difference between the mean scores on the five flight phases.

The flight phases pairwise comparisons considering each physiological index were inspected, using Bonferroni correction, in order to identify which mean scores were statistically different from each other.

A one-way ANOVA with repeated measures showed a significant main effect of flight phases on SI ( $F_{(4,132)} = 25.261, p < 0.001$ ). Significant differences in mean SI were found between STD and all the other flight phases (all differences were significant at  $p < 0.001$ ), between ENR and MNV ( $p = 0.007$ ), and between ENR and LDG ( $p = 0.001$ ).

The results of one-way ANOVA with repeated measures revealed a significant main effect of flight phases on SDNN ( $F_{(2,804,92.524)} = 41.995, p < 0.001$ ). A statistical difference for mean SDNN scores was obtained between STD and all the other flight phases (all differences were significant at  $p < 0.001$ ), between ENR and TOF ( $p = .002$ ), between ENR and MNV as well as between ENR and LDG ( $p < 0.001$ ).

TABLE 5. ANOVA with repeated measures on flight phases.

Index	STD		ENR		TOF		MNV		LDG		ANOVA results	
	M	SD	M	SD	M	SD	M	SD	M	SD	F	$\eta^2$
SI	12.77	4.58	19.73	11.09	22.51	9.46	24.88	12.29	25.56	12.29	25.26	<b>0.331</b>
SDNN	41.70	15.01	29.29	15.98	21.37	12.06	20.72	13.54	20.48	13.98	41.99	<b>0.452</b>
RMSSD	30.54	17.64	21.52	19.12	16.35	15.12	16.00	16.91	16.38	16.48	12.80	<b>0.223</b>
HRV <sub>ti</sub>	9.51	3.28	6.46	2.83	4.62	2.22	4.73	2.66	4.43	2.32	60.16	<b>0.562</b>
TINN	224.03	78.27	164.88	89.51	134.94	78.82	125.47	87.18	126.53	84.23	25.44	<b>0.333</b>
VLF	86.61	70.05	53.55	64.94	34.03	46.78	35.51	42.89	30.16	38.17	9.19	<b>0.194</b>
LF	1206.1	807.10	701.25	639.17	316.30	335.28	337.01	376.39	320.83	379.88	30.82	<b>0.430</b>
HF	520.41	588.79	333.41	611.54	142.37	276.89	157.11	316.73	134.94	228.70	8.44	<b>0.174</b>
TOT <sub>pow</sub>	1816.7	1319.9	1088.4	1186.3	493.1	599.3	529.8	672.4	486.1	594.7	23.94	<b>0.367</b>
LF/HF	4.16	2.78	6.06	6.37	5.14	4.09	6.52	6.31	5.71	4.57	1.93	0.045
SD1	21.63	12.50	15.24	13.54	11.58	10.71	11.33	11.97	11.70	11.67	12.81	<b>0.204</b>
SD2	54.46	18.43	37.80	19.54	27.28	14.38	26.49	15.94	25.88	17.03	53.04	<b>0.501</b>

Note: Except for LF/HF ( $p = 0.108$ ), all effects are significant at  $< 0.001$

TABLE 4. ANOVA between age groups.

Index	Age	M	SD	F	p	$\eta^2$
SI	$\leq 30$	12.414	4.886	0.18	0.67	0.01
	$> 30$	13.091	4.42			
SDNN	$\leq 30$	44.937	16.874	1.43	0.24	0.04
	$> 30$	38.816	12.946			
RMSSD	$\leq 30$	33.638	18.401	0.93	0.34	0.03
	$> 30$	27.791	16.978			
HRV <sub>ti</sub>	$\leq 30$	10.192	3.624	1.32	0.26	0.04
	$> 30$	8.902	2.916			
TINN	$\leq 30$	233.188	74.534	0.41	0.53	0.01
	$> 30$	215.889	82.718			
VLF	$\leq 30$	93.279	74.299	0.27	0.61	0.01
	$> 30$	80.688	67.651			
LF	$\leq 30$	1410.31	949.756	1.99	0.17	0.06
	$> 30$	1024.67	628.088			
HF	$\leq 30$	739.555	745.246	4.65	0.04	0.13
	$> 30$	325.61	312.292			
TOT <sub>pow</sub>	$\leq 30$	2243.74	1640.5	3.45	0.07	0.1
	$> 30$	1431.41	823.559			
LF/HF	$\leq 30$	2.763	1.513	9.71	0	0.23
	$> 30$	5.407	3.075			
SD1	$\leq 30$	23.82	13.035	0.93	0.34	0.03
	$> 30$	19.679	12.025			
SD2	$\leq 30$	58.489	21.224	1.47	0.24	0.04
	$> 30$	50.875	15.252			

Note: significant differences are in bold.

Regarding RMSSD, the findings indicated a significant main effect ( $F_{(2,899,92.695)} = 12.803, p < 0.001$ ). Statistical differences were found between STD and ENR ( $p = 0.007$ ), between STD and TOF ( $p < 0.001$ ), between STD and MNV ( $p < 0.001$ ), and between STD and LDG ( $p = 0.001$ ). Another significant difference in RMSSD mean scores was estimated between ENR and MNV ( $p = 0.033$ ).

The findings showed a significant main effect for HRV triangular index ( $F_{(2,307,76.147)} = 60.160, p < 0.001$ ). Significant differences were estimated between STD and all the other flight phases, between ENR and TOF, between ENR and MNV, between ENR and LDG (all differences were significant at  $p < 0.001$ ).

The results of the one way ANOVA reported a significant main effect of flight phases for mean TINN ( $F_{(2,853,94.148)} = 25.443, p < 0.001$ ). Significant differences were found between STD and all the other flight phases (all differences were significant at  $p < 0.001$ ). In addition, a significant difference was observed between ENR and MNV ( $p = 0.002$ ) and between ENR and LDG ( $p = 0.003$ ).

**TABLE 6.** Boolean matrix of significant difference between flight phases (for the considered index 1 = significant difference and 0 = not significant difference).

Index	STD-ENR	STD-TOF	STD-MNV	STD-LDG	ENR-TOF	ENR-MNV	ENR-LDG	TOF-MNV	TOF-LDG	MNV-LDG
SDNN	1	1	1	1	1	1	1	0	0	0
SD2	1	1	1	1	1	1	1	0	0	0
LF	1	1	1	1	1	1	1	0	0	0
$TOT_{pow}$	1	1	1	1	1	1	1	0	0	0
$HRV_{ti}$	1	1	1	1	1	1	1	0	0	0
SI	1	1	1	1	0	1	1	0	0	0
TINN	1	1	1	1	0	1	1	0	0	0
RMSSD	1	1	1	1	0	1	0	0	0	0
SD1	1	1	1	1	0	1	0	0	0	0
VLF	1	1	1	0	0	0	0	0	0	0
HF	1	1	1	0	0	0	0	0	0	0
LH/HF	0	0	0	0	0	0	0	0	0	0

VLF power resulted in a ( $F_{(2,803,92,504)} = 9.192, p = .000$ ), statistical differences were estimated between STD and TOF ( $p = 0.003$ ), between STD and MND ( $p = 0.004$ ), and between STD and LDG ( $p < 0.001$ ).

Regarding the LF power, the results of one-way ANOVA with repeated measurements showed a mean significant effect ( $F_{(2,300,79,909)} = 30.820, p < 0.001$ ). The pairwise STD and ENR reported statistical differences ( $p = .007$ ), whereas the pairwise STD with TOF or MNV, or LDG reported values of significance  $p < 0.001$ . Other significant differences were found for the LF index between ENR and TOF with ( $p < 0.001$ ), and between ENR and MNV or LDG ( $p = 0.003$ ).

Similarly for HF power, the tests revealed a significant mean effect ( $F_{(2,310,76,227)} = 8.440, p < 0.001$ ), with significant differences between STD and TOF ( $p = 0.007$ ), between STD and MNV ( $p < 0.001$ ), and between STD and LDG ( $p = 0.003$ ).

With respect to  $TOT_{pow}$ , a value of ( $F_{(2,352,76,606)} = 23.940, p < 0.001$ ) was estimated with significant differences between STD and TOF or MNV, or LDG ( $p < 0.001$ ). Other significant differences were found between ENR and TOF ( $p = 0.005$ ), between ENR and MNV ( $p = 0.007$ ), and between ENR and LDG ( $p = 0.003$ ).

Concerning SD1, the results of one way ANOVA with repeated measures report a significant mean effect ( $F_{(2,989,95,645)} = 12.813, pp < 0.001$ ). Significant differences are shown between STD and ENR ( $p = 0.007$ ), between STD and TOF or MNV results ( $p < 0.001$ ), and between STD and LDG ( $p = 0.001$ ). Another significant difference was also estimated in the comparison between ENR and MNV ( $p = 0.032$ ), whereas no statistical differences were found with TOF or LDG.

Lastly, the results of one way ANOVA with repeated measures indicated a significant mean effect for SD2 ( $F_{(2,849,94,028)} = 53.041, p < 0.001$ ). Statistical differences were reported between STD and all the other flight phases, and between ENR and TOF, ENR and MNV, and between ENR and LDG (all differences are significant at  $p < 0.001$ ).

Table 6 shows a synthetic resume scheme of comparisons between each flight phase for each index, in which statistically significant differences are indicated by the boolean marker “1”.

Lastly, the results of the mixed designed ANOVA test reported no significant interaction between age and flight phases, Wilk’s  $\lambda = 0.670, F_{(48,452,735)} = 1.035, p = 0.413$ , partial  $\eta^2 = 0.095$ . This means that changes in mean scores on the physiological measures over the different flight phases were equivalent across the two age groups.

**E. FLIGHT PHASES COMPARISON**

Figure 2 shows the notched boxplots normalized to the reference baseline of the STD to produce a visual interpretation of the values of the HRV indexes obtained over the different flight phases. The results described in the previous paragraph have highlighted the significant differences obtained between the various phases; however, to identify whether the workload levels were higher or lower, the trend of each index was analyzed over the flight phases. In Figure 2, a red line between the values of the medians is shown. Analyzing the indexes’ tendencies, it is possible to view how all indexes have a well-defined trend between the baseline and the ENR phase and with the three phases of TOF, MNV, and LDG, which do not exhibit noticeable variations between them. This suggests that the workload levels are comparable for the Category C terminal phases of TOF and LDG and a constrained maneuver such as the steady turn.

This holds for all the indexes with the exception of LF/HF because in the pairwise comparisons no significant differences have emerged using one-way ANOVA with repeated measures. In Figure 2 this effect is noticeable as the fluctuation of the index between ENR and MNV compared to TOF and LDG. Concerning SI, which shows an increasing trend compared to the other indexes, it should be recognized that an increase in SI is associated with an increase in workload levels; hence, the results agree with the other indexes.

Considering the first five indexes in Table 6, namely SDNN, SD2, LF,  $TOT_{pow}$ , and  $HRV_{ti}$ , Figure 2 presents a clear trend with respect to the baseline. For these indexes, the whiskers of dispersion are more contained than those for the other indexes that present at least one flight phase with a much greater variance than the other or with skewed data. Concerning LF and  $TOT_{pow}$ , the ENR whiskers have a greater extension than the other phases; this effect is also found in all the other frequency bands.

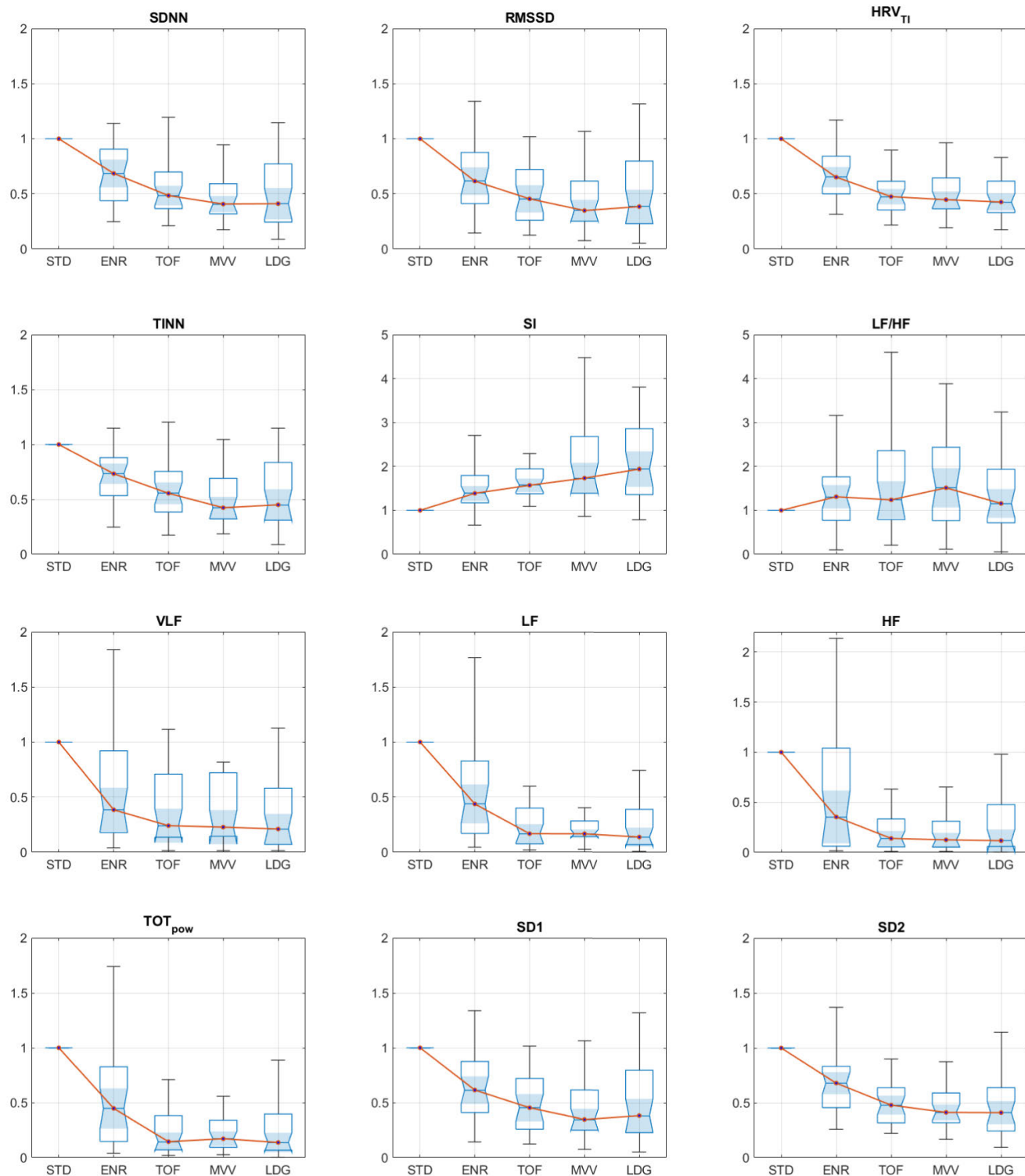


FIGURE 2. Notched boxplot of flight phases.

### F. WORKLOAD ASSESSMENT

A preliminary method of workload levels estimation, based on specific considerations for the first five indexes listed in Table 6, is detailed in this section. These indexes show similar results in the pairwise comparison of the flight phases.

In particular, based on what was mentioned in the previous section, the TOF, LDG, and MNV flight phases can be associated with a high workload level. With the purpose of using three different workload levels (low, medium and high), the data of the three high-workload phases were merged at a



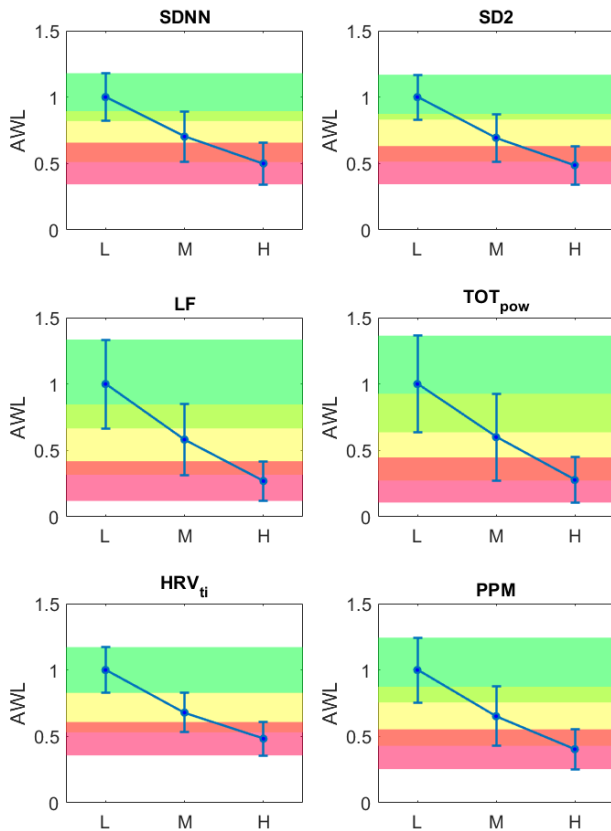


FIGURE 3. Assessment of allowable workload levels.

high level. See Figure 3 that shows the trends of the previously selected indexes. The y-axis is named AWL- Allowable WorkLoad, meant by authors as the percentage amount of workload that the subjects are still able to use. Thus, a value equal to one (i.e., 100%) is associated with the lowest level of workload. The deviation with respect to the mean value is computed as 0.5 times the standard deviation of the acquired sample. Three different bands are drawn with color gradation, and the highest workload assessment zone is colored in red. Except for the green and yellow bands in the  $HRV_{ti}$  index, all others bands show an overlapped transition band from lower to higher workload level.

A sixth further index is introduced in Figure 3, built by combining the previous five selected indexes; this index is named PPM - Pilot Performance Monitoring. The PPM is calculated starting from the normalized values for the baseline rest time of the five indexes considered and summing the contributions of each index for the corresponding difficulty level. The overlap thresholds shown in Figure 3 for PPM were obtained considering a standard deviation of 0.5, which results in central overlapping bands that differentiate the AWL thresholds more systematically than the results obtained by the individual indexes.

A multivariate analysis of variance was conducted to test the overall significance of the three workload levels among the five selected measures. Setting a level of  $\alpha = 0.05$ , a significant main effect was estimated, Wilks'  $\lambda = 0.105$ ,

TABLE 7. Analysis results for the three AWL levels.

Measure	Workload level (M $\pm$ SD)			F	Partial $\eta^2$
	Low	Medium	High		
SDNN	41.70 $\pm$ 15.01	29.29 $\pm$ 15.98	20.85 $\pm$ 11.71	63.99	0.66
$HRV_{ti}$	9.51 $\pm$ 3.28	6.46 $\pm$ 2.83	4.58 $\pm$ 2.23	78.76	0.71
LF	1206.15 $\pm$ 807.11	701.25 $\pm$ 639.16	324.72 $\pm$ 313.13	29.54	0.47
$TOT_{pow}$	1816.68 $\pm$ 1319.87	1088.45 $\pm$ 1186.35	502.98 $\pm$ 522.56	22.16	0.41
SD2	54.46 $\pm$ 18.43	37.80 $\pm$ 19.54	26.55 $\pm$ 14	76.25	0.70
PPM	3125.50 $\pm$ 2126.85	1863.25 $\pm$ 1818.41	879.69 $\pm$ 853.57	26.53	0.45

$F_{(10,24)} = 20.446$ ,  $p < 0.001$ , partial  $\eta^2 = 0.895$ , meaning that the combination of the five selected physiological parameters provides a useful discrimination of the three different workload levels. In addition, the results showed that there were statistical differences among the three different workload levels for each measure (Table 7).

Wilks'  $\lambda$  value in a multivariate test does not provide explicit information on how the indexes contribute to differentiating the workload levels for the analyzed case study. Wilks'  $\lambda$  give only the knowledge that there is a significant effect on at least one of the dependent variables taken into consideration.

Thus, repeated contrasts were chosen; each contrast compared the index's value at one level to the index's value at the following level, and significance values were computed to verify that all the indexes can recognize the differences.

In the low-medium comparison (L-M), the two contrasts listed in Table 8,  $TOT_{pow}$  and LF, had significance values greater than or equal to 0.001. The cause of reduced significant effect in these two comparisons is detectable from the analysis of boxplot data described in the previous paragraph flight phases comparison, in which  $TOT_{pow}$  and LF reported higher whiskers due to the variability of acquired values in the ENR flight phase.

Nevertheless, by identifying possible outliers in the sample for each level using the multivariate Mahalanobis distance test, four pilots were found in at least one of the levels with remarkably high distance values. Excluding these pilots and reanalyzing the sample, based on 30 pilots, all indexes exhibited significant differences in the comparison between levels. The results for the sample of 30 pilots are shown in Table 9. Descriptives, F values, and effect size (Partial  $\eta^2$ ) were reported for the analyzed groups.

## V. DISCUSSION

Concerning the influence of sex (data reported in Table 3), except for the SI, none of the investigated physiological indexes showed statistical differences between males and females. However, a recent meta-analysis [64] systematically reported the relationship between HRV indexes and sex. For instance, analysis of time-domain HRV indexes showed a more decreased mean score on SDNN in females than in males, whereas analysis of RMSSD yielded no

**TABLE 8.** Repeated contrasts analysis group with outliers.

Group of 34		F	<i>p</i>	Partial $\eta^2$
SDNN	L-M	40,010	< 0.001*	0,548
	M-H	39,159	< 0.001*	0,543
HRV <sub>ti</sub>	L-M	49,656	< 0.001*	0,601
	M-H	59,920	< 0.001*	0,645
LF	L-M	13,855	,001	0,296
	M-H	21,843	< 0.001*	0,398
TOT <sub>pow</sub>	L-M	8,723	,006	0,209
	M-H	17,803	< 0.001*	0,350
SD2	L-M	47,716	< 0.001*	0,591
	M-H	44,474	< 0.001*	0,574

**TABLE 9.** Repeated contrasts analysis group without outliers.

Group of 30		F	<i>p</i>	Partial $\eta^2$
SDNN	L-M	60,685	< 0.001*	0,677
	M-H	30,843	< 0.001*	0,515
HRV <sub>ti</sub>	L-M	52,685	< 0.001*	0,645
	M-H	59,012	< 0.001*	0,670
LF	L-M	16,770	< 0.001*	0,366
	M-H	19,452	< 0.001*	0,401
TOT <sub>pow</sub>	L-M	21,078	< 0.001*	0,421
	M-H	19,972	< 0.001*	0,408
SD2	L-M	58,879	< 0.001*	0,670
	M-H	39,187	< 0.001*	0,575

significant or inconsistent findings. Furthermore, regarding the frequency-domain HRV indexes, females compared to males reported lower scores on VLF, LF, LF/HF ratio, TOT<sub>pow</sub>, and higher HF [64]. These differences between the present results and the literature may be partly due to the participants' characteristics of the examined studies, which may have affected HRV indexes scores between the two sexes. To include these effects in the workload assessment, it is important to investigate a sample of pilots that includes more females than those involved in the present study. The authors suggest considering these findings with some caution since they may be influenced by the high disproportion of the number of male and female participants. However, it is worth noting that the percentage of females involved in the experiment is in line with the current civil pilots population [65], [66].

Regarding the influence of age, in the present study, only HF and LF/HF showed significant differences between the two age groups, see Table 4. HF was negatively associated with age, suggesting a reduced autonomic tone of the heart with an increase in age. On the other hand, positive associations were estimated with LF/HF, indicating an increment of sympathetic activity with an increase in age. These results agree with previous empirical works [67], [68], which found associations characterized by similar trend. Nevertheless, the results of the current study do not support the existence of any influence of age on the other physiological indexes. There is a possible explanation for these findings. For instance, in some studies, age is negatively related to SDNN, to RMSSD, and TOT<sub>pow</sub> [67]–[69], but it is important to outline that in the aforementioned studies, the participants' age groups were very larger and distinguished (aged from 16 to 60,

from 40 to 100, and from 10 to 80, respectively). In contrast, the present study sample was restricted to only adult individuals (min age = 23 years, max age = 52 years). It is likely that, considering different age groups, significant results would have been found. Nevertheless, it is worth noting that the age of pilots involved in the present study denotes a close proneness to a realistic training group.

Concerning the trend of the HRV indexes over the flight phases, some general aspects can be highlighted. In particular, it is possible to note how an increase in the workload affected each index. Figure 2 showed for the SI index a different response from the other indexes; in fact, the value for the in-flight phases was greater than the baseline. This result agrees with the pilots' physiological response since, according to Baevsky *et al.* [45] SI increases as the workload increases. By analyzing the time domain indexes SDNN and RMSSD, it can be observed that they had a decrement related to an increase in workload level; comparable results were found by Cinaz *et al.* [70] and Orsila *et al.* [43]. The differences in the mean values were estimated in the present study with a significant difference for SDNN, but even to the decrease in RMSSD according to Cinaz *et al.* [71] and Fallahi *et al.* [52] should be related to an increase in mental workload. Similarly, a decrease in the geometric index is related to an increase in workload [70], [71]. More specifically, a decrease in the HRV<sub>ti</sub> index is associated with an increase in the workload levels. This resulted in a significant difference in the pairwise comparisons of flight phases according to Mansikka *et al.* [4], [5]; the TINN index resulted in a weak significant difference in the pairwise comparisons of flight phases, which is in accordance with Cinaz *et al.* [71]. Regarding the frequency domain, the results are largely in agreement with the reference literature. In particular, the obtained results showed that a decrease in both LF and HF bands reflects an increase in the workload. Regarding these two frequency bands also Bonner and Wilson [28] found a significant decrease as the workload increased.

Separate considerations must be considered for the LF/HF ratio. The results of the present work show that all the flight phases compared with the baseline have an average level that is higher than the baseline itself. This means that the workload level during rest time is lower than during flight phases; however, for the pilots and flight phases analyzed, the LF/HF ratio is inadequate to differentiate between medium and high workload levels. It is worth mentioning here that the results reported in the scientific literature about the relationships between the LF/HF ratio and workload levels are not always in agreement. In some previous works [30], [51]–[53] it was reported that an increase in the LF/HF ratio is associated with a workload increase. Conversely, the opposite trend was reported by Hsu *et al.* [54]. The present study as well does not reveal a distinct relationships between WL and flight phases based on the examined sample.

The other two indexes, VLF and TOT<sub>pow</sub>, examined in the frequency domain, reported a decrease associated with an increase in workload, and similar results were found by

Heine et al. [72] and Miyake et al. [73]. More in particular, the decrease in VLF results is sensitive enough to differentiate the maneuvers associated with high mental workloads, this is in accordance with Hsu et al. [54];  $TOT_{pow}$  presents a narrower variance and a significant difference in pairwise comparison of flight phases, and for this reason, it is an appropriate index for the workload level inspection.

Finally, the non linear indexes computed on Poincaré plot, SD1, and SD2 showed a reduction in the median values that should be related to a workload increase as found by Acharya et al. [74]. Both SD1 and SD2 reduced as the workload increased, so the Poincaré ellipse contraction can be related to variations in the workload levels.

As a further result, all the considered HRV indexes' analysis proved that takeoff and landing maneuvers, belonging to Category C handling qualities classification, have similar workload levels to a double steady turn maneuver; therefore, it is possible to consider these three maneuvers as a group to high demand workload.

Finally, with the aim of differentiating among low, medium and high workload levels, the indexes SDNN, SD2, LF,  $TOT_{pow}$ , and  $HRV_{i}$  were selected in the present work because they showed a statistically significant decrease with workload increase and because they allowed the estimation of the three different pilots AWL bands. In addition to these findings, an index named PPM - Pilot Performance Monitoring was introduced as the sum of the five HRV indexes. For the considered case study, the PPM seems to assess more consistently the allowable workload level. These results can be considered a step forward in the identification of the workload level that pilots can experience during flight or training since, to the best knowledge of the authors, this is the first study that produces consistency in the differentiation among low, medium, and high workload levels using the time, frequencies, and non-linear HRV domains of analysis.

## VI. LIMITATIONS AND FUTURE DIRECTIONS

The main limitation of the study could be seen in the low number of participants which could limit the generalization of the findings beyond the study sample. Nevertheless, it is worth noting that the sample size is similar to the ones reported in other previous works about flight simulation activities and that the results are in line with the literature [4], [8], [52], [54].

A second limitation concerns the inclusion of standard maneuvers, excluding from analysis any non-standard maneuvers, which may be likely associated with higher workload levels. From this perspective, it will be necessary to verify what occurs with much more challenging maneuvers, such as the final approach with low visibility or one engine inoperative.

Another limitation may be that the experiments were carried out using a ground-based simulator instead of real flying aircraft. However, it is clear that using a ground-based simulator is consistent with first steps research studies to lay down the basis of future experiments to be carried out using

real aircraft simulator systems such as the Calspan's Learjet Simulators.

Additional studies should also be addressed at taking into account shorter time intervals than five minutes to investigate whether the same conclusions can be drawn also in a different time frame.

For future testing, the optimal and recommendable choice is the detection of multiple HRV parameters, since a single indicator may not be able to fully describe the workload. From this point of view, the information coming from different indexes can converge in the same direction, providing a more detailed description of the investigated workload assessment.

## VII. CONCLUSION

This study can be considered a contribution to workload assessment using physiological measures. Specifically, considering time-domain, frequency-domain, and non-linear HRV indexes in a unique experimental design, these findings offer a more comprehensive overview of the different HRV indexes for workload evaluation. These results support the importance of parameters based on HRV measurement for assessing changes in workload, also suggesting the development of non-obtrusive devices that can assess workload in real-time.

## REFERENCES

- [1] D. L. Peahl, D. R. Sandlin, and G. Ncc, "A survey of handling qualities criteria and their applications," High Perform. Aircr. Principal Investigators Faysal Kolkailah, Aeronaut. Eng. Dept., San Luis Obispo, CA, USA, Final Rep. NASA-CR-179654, 1986.
- [2] G. E. Cooper and R. P. Harper, "The use of pilot ratings in the evaluation of aircraft handling qualities," Advisory Group Aerosp. Res. Develop., Neuilly-Sur-Seine, France, Tech. Rep. AGARD-567, 1969, pp. 1–56.
- [3] N. Dahlstrom and S. Nahlinger, "Mental workload in aircraft and simulator during basic civil aviation training," *Int. J. Aviation Psychol.*, vol. 19, no. 4, pp. 309–325, 2009.
- [4] H. Mansikka, K. Virtanen, D. Harris, and P. Simola, "Fighter pilots' heart rate, heart rate variation and performance during an instrument flight rules proficiency test," *Appl. Ergonom.*, vol. 56, pp. 213–219, Sep. 2016.
- [5] H. Mansikka, P. Simola, K. Virtanen, D. Harris, and L. Oksama, "Fighter pilots' heart rate, heart rate variation and performance during instrument approaches," *Ergonomics*, vol. 59, no. 10, pp. 1344–1352, 2016.
- [6] P. H. Lehmann, M. Jones, and M. Höfinger, "Impact of turbulence and degraded visual environment on pilot workload," *CEAS Aeronaut. J.*, vol. 8, no. 3, pp. 413–428, Sep. 2017.
- [7] K. Mohanavelu, S. Poonguzhali, D. Ravi, P. K. Singh, M. Mahajabin, K. Ramachandran, U. K. Singh, and S. Jayaraman, "Cognitive workload analysis of fighter aircraft pilots in flight simulator environment," *Defence Sci. J.*, vol. 70, no. 2, pp. 131–139, Mar. 2020.
- [8] J. P. Fuentes-García, V. J. Clemente-Suárez, M. A. Marazuela-Martínez, J. F. Tornero-Aguilera, and S. Villafaina, "Impact of real and simulated flights on psychophysiological response of military pilots," *Int. J. Environ. Res. Public Health*, vol. 18, no. 2, p. 787, 2021.
- [9] D. H. Klyde, A. Lampton, D. G. Mitchell, C. Berka, and M. Rhinehart, "A new approach to aircraft handling qualities prediction," in *AIAA Scitech Forum*, 2021, pp. 1–15.
- [10] A. Alaimo, A. Esposito, and C. Orlando, "Cockpit pilot warning system: A preliminary study," in *Proc. IEEE 4th Int. Forum Res. Technol. Soc. Ind. (RTSI)*, Sep. 2018, pp. 1–4.
- [11] H. Ji, R. Chen, L. Lu, and M. D. White, "Pilot workload investigation for rotorcraft operation in low-altitude atmospheric turbulence," *Aerosp. Sci. Technol.*, vol. 111, Apr. 2021, Art. no. 106567.

- [12] E. N. Bachelder and M. Godfroy-Cooper, "Pilot workload estimation: Synthesis of spectral requirements analysis and weber's law," in *Proc. AIAA Scitech Forum*, 2019, p. 1228.
- [13] G. G. De la Torre, M. A. Ramallo, and E. Cervantes, "Workload perception in drone flight training simulators," *Comput. Hum. Behav.*, vol. 64, pp. 449–454, Nov. 2016.
- [14] G. Matthews, J. De Winter, and P. A. Hancock, "What do subjective workload scales really measure? Operational and representational solutions to divergence of workload measures," *Theor. Issues Ergonom. Sci.*, vol. 21, no. 4, pp. 369–396, Jul. 2020.
- [15] H. Mansikka, K. Virtanen, and D. Harris, "Comparison of NASA-TLX scale, modified Cooper–Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks," *Ergonomics*, vol. 62, no. 2, pp. 246–254, 2019.
- [16] A. Alaimo, A. Esposito, C. Orlando, and A. Simoncini, "Aircraft pilots workload analysis: Heart rate variability objective measures and NASA-task load index subjective evaluation," *Aerospace*, vol. 7, no. 9, p. 137, Sep. 2020.
- [17] E. Buccelletti, E. Gilardi, E. Scaini, L. Galiuto, R. Persiani, A. Biondi, F. Basile, and N. G. Silveri, "Heart rate variability and myocardial infarction: Systematic literature review and meta-analysis," *Eur. Rev. Med. Pharmacol. Sci.*, vol. 13, no. 4, pp. 299–307, Jul./Aug. 2009.
- [18] H. ChuDuc, K. NguyenPhan, and D. NguyenViet, "A review of heart rate variability and its applications," *APCBEE Proc.*, vol. 7, pp. 80–85, Jan. 2013.
- [19] H. G. Kim, E. J. Cheon, D. S. Bai, Y. H. Lee, and B. H. Koo, "Stress and heart rate variability: A meta-analysis and review of the literature," *Psychiatry Invest.*, vol. 15, no. 3, pp. 235–245, 2018.
- [20] J. Kranjec, S. Beguš, G. Geršak, and J. Drnovšek, "Non-contact heart rate and heart rate variability measurements: A review," *Biomed. Signal Process. Control*, vol. 13, pp. 102–112, Sep. 2014.
- [21] R. Shashikant and P. Chetankumar, *A Review on Impact Application of Heart Rate Variability (HRV)*. Cham, Switzerland: Springer, 2020.
- [22] G. F. Wilson, "An analysis of mental workload in pilots during flight using multiple psychophysiological measures," *Int. J. Aviation Psychol.*, vol. 12, no. 1, pp. 3–18, Jan. 2002.
- [23] K. Alghoul, S. Alharthi, A. El Saddik, and H. Al Osman, "Heart rate variability extraction from videos signals: ICA vs. EVM comparison," *IEEE Access*, vol. 5, pp. 4711–4719, 2017.
- [24] T. Iqbal, P. Redon-Lurbe, A. J. Simpkin, A. Elahi, S. Ganly, W. Wijns, and A. Shahzad, "A sensitivity analysis of biophysiological responses of stress for wearable sensors in connected health," *IEEE Access*, vol. 9, pp. 93567–93579, 2021.
- [25] B. Cain, "A review of the mental workload literature," Defence Res. Develop., Toronto, ON, Canada, Tech. Rep. RTO-TR-HFM-121-Part-II-4-1, 2007, p. 4.
- [26] R. L. Charles and J. Nixon, "Measuring mental workload using physiological measures: A systematic review," *Appl. Ergonom.*, vol. 74, pp. 221–232, Jan. 2019.
- [27] D. Tao, H. Tan, H. Wang, X. Zhang, X. Qu, and T. Zhang, "A systematic review of physiological measures of mental workload," *Int. J. Environ. Res. Public Health*, vol. 16, no. 15, p. 2716, 2019.
- [28] M. A. Bonner and G. F. Wilson, "Heart rate measures of flight test and evaluation," *Int. J. Aviation Psychol.*, vol. 2014, pp. 37–41, Oct. 2009.
- [29] P. H. Gander, T. L. Signal, M. J. van den Berg, H. M. Mulrine, S. M. Jay, and C. Jim Mangie, "In-flight sleep, pilot fatigue and psychomotor vigilance task performance on ultra-long range versus long range flights," *J. Sleep Res.*, vol. 22, no. 6, pp. 697–706, Dec. 2013.
- [30] A. Alaimo, A. Esposito, C. Orlando, and G. Tesoriere, "A pilot mental workload case study in a full flight simulator," *Aerotecnica Missili Spazio*, vol. 97, no. 1, pp. 27–33, Jan. 2018.
- [31] S. Gedam and S. Paul, "A review on mental stress detection using wearable sensors and machine learning techniques," *IEEE Access*, vol. 9, pp. 84045–84066, 2021.
- [32] A. Alaimo, A. Esposito, A. Milazzo, and C. Orlando, "An aircraft pilot workload sensing system," in *Proc. Eur. Workshop Struct. Health Monit.* (Lecture Notes in Civil Engineering), vol. 127. Cham, Switzerland: Springer, 2021, pp. 883–892.
- [33] P. G. Jorna, "Heart rate and workload variations in actual and simulated flight," *Ergonomics*, vol. 36, no. 9, pp. 1043–1054, 1993.
- [34] A. J. Tattersall and G. R. J. Hockey, "Level of operator control and changes in heart rate variability during simulated flight maintenance," *Hum. Factors*, vol. 37, no. 4, pp. 682–698, Dec. 1995.
- [35] T. M. Lahtinen, J. P. Koskelo, T. Laitinen, and T. K. Leino, "Heart rate and performance during combat missions in a flight simulator," *Aviation, Space, Environ. Med.*, vol. 78, no. 4, pp. 387–391, 2007.
- [36] F. Sauvet, J. C. Jouanin, C. Langrume, P. Van Beers, Y. Papelier, and C. Dussault, "Heart rate variability in novice pilots during and after a multi-leg cross-country flight," *Aviation, Space, Environ. Med.*, vol. 80, no. 10, pp. 862–869, Oct. 2009.
- [37] E.-R. Saus, B. H. Johnsen, J. Eid, and J. F. Thayer, "Who benefits from simulator training: Personality and heart rate variability in relation to situation awareness during navigation training," *Comput. Hum. Behav.*, vol. 28, no. 4, pp. 1262–1268, Jul. 2012.
- [38] X. Cao, P. MacNaughton, L. Cadet, J. Cedeno-Laurent, S. Flanigan, J. Vallarino, D. Donnelly-McLay, D. Christiani, J. Spengler, and J. Allen, "Heart rate variability and performance of commercial airline pilots during flight simulations," *Int. J. Environ. Res. Public Health*, vol. 16, no. 2, p. 237, Jan. 2019.
- [39] K. Čosić, M. Šarlija, V. Ivkovic, Q. Zhang, G. Strangman, and S. Popović, "Stress resilience assessment based on physiological features in selection of air traffic controllers," *IEEE Access*, vol. 7, pp. 41989–42005, 2019.
- [40] P. Kutilek, P. Volf, K. Sedova, J. Hejda, V. Krivanek, M. Stehlík, K. Rusnakova, S. Kozlova, and M. Braunova, "Heart rate variability during fighter pilot training preliminary study," in *Proc. Int. Conf. Mil. Technol. (ICMT)*, May 2019, pp. 1–5.
- [41] A. Alaimo, A. Esposito, and C. Orlando, "Pilot's mental workload nonlinear correlation among objective and subjective measurements," in *Proc. J. Phys., Conf.*, 2021, vol. 1786, no. 1, Art. no. 012011.
- [42] Task Force of the European Society Organization Electrophysiology, "Heart rate variability," *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar. 1996.
- [43] R. Orsila, M. Virtanen, T. Luukkaala, M. Tarvainen, P. Karjalainen, J. Viik, and M. Savinainen, "Perceived mental stress and reactions in heart rate variability—A pilot study among employees of an electronics company," *Int. J. Occupational Saf. Ergonom.*, vol. 14, no. 3, pp. 275–283, 2008.
- [44] R. M. Baevsky, "Noninvasive methods in space cardiology," *J. Cardiovascular Diagnosis Procedures*, vol. 14, no. 3, pp. 161–171, 1997.
- [45] R. M. Baevsky, V. M. Petrov, and A. G. Chernikova, "Regulation of autonomic nervous system in space and magnetic storms," *Adv. Space Res.*, vol. 22, no. 2, pp. 227–234, Jan. 1998.
- [46] M. P. Tarvainen, J. P. Niskanen, J. A. Lipponen, P. O. Ranta-aho, and P. A. Karjalainen, "Kubios HRV—heart rate variability analysis software," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 210–220, 2014.
- [47] A. Malliani, M. Pagani, F. Lombardi, and S. Cerutti, "Cardiovascular neural regulation explored in the frequency domain," *Circulation*, vol. 84, no. 2, pp. 482–492, 1991.
- [48] J. Sztajzel, "Heart rate variability: A noninvasive electrocardiographic method to measure the autonomic nervous system," *Swiss Med. Weekly*, vol. 134, nos. 35–36, pp. 514–522, 2004.
- [49] R. E. Kleiger, P. K. Stein, and J. T. Bigger, "Heart rate variability: Measurement and clinical utility," *Ann. Noninvasive Electrocardiol.*, vol. 10, no. 1, pp. 88–101, Jan. 2005.
- [50] F. Shaffer, R. McCraty, and C. L. Zerr, "A healthy heart is not a metronome: An integrative review of the heart's anatomy and heart rate variability," *Frontiers Psychol.*, vol. 5, pp. 1–19, Sep. 2014.
- [51] A. Bilan, A. Witzak, R. Palusiński, W. Myśliński, and J. Hanzlik, "Circadian rhythm of spectral indices of heart rate variability in healthy subjects," *J. Electrocardiol.*, vol. 38, no. 3, pp. 239–243, Jul. 2005.
- [52] M. Fallahi, M. Motamedzade, R. Heidarimoghadam, A. R. Soltanian, and S. Miyake, "Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study," *Appl. Ergonom.*, vol. 52, pp. 95–103, Jan. 2016.
- [53] F. W. Skibniewski, L. Dziuda, P. M. Baran, M. K. Krej, S. Guzowski, M. A. Piotrowski, and O. E. Trusczyński, "Preliminary results of the LF/HF ratio as an indicator for estimating difficulty level of flight tasks," *Aerosp. Med. Human Perform.*, vol. 86, no. 6, pp. 518–523, Jun. 2015.
- [54] B.-W. Hsu, M.-J.-J. Wang, C.-Y. Chen, and F. Chen, "Effective indices for monitoring mental workload while performing multiple tasks," *Perceptual Motor Skills*, vol. 121, no. 1, pp. 94–117, Aug. 2015.
- [55] M. Valderrama, V. Navarro, and M. Le Van Quyen, "Heart rate variability as measurement of heart-brain interactions," *Epilepsies*, vol. 22, no. 3, pp. 194–200, Jul. 2010.
- [56] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers Public Health*, vol. 5, pp. 1–17, Sep. 2017.

- [57] P. W. Kamen, H. Krum, and A. M. Tonkin, "Poincaré plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans," *Clin. Sci.*, vol. 91, no. 2, pp. 201–208, 1996.
- [58] M. Brennan, M. Palaniswami, and P. Kamen, "Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 283, no. 5, pp. 1873–1886, 2002.
- [59] S. K. Ghatak and S. Aditya, "Poincaré parameters and principal component analysis of Heart rate variability of subjects with health disorder," 2018, *arXiv:1802.10289*.
- [60] Department of Defense, *Flying Qualities of Piloted Aircraft*, Standard Mil-Std-1797a, 2004.
- [61] M. A. García-González, A. Argelagós-palau, M. Fernández-chimeno, J. Ramos-castro, and P. D. Catalunya, "A comparison of heartbeat detectors for the seismocardiogram," in *Proc. Comput. Cardiol.*, Sep. 2013, pp. 461–464.
- [62] M. Vollmer, D. Bläsing, and L. Kaderali, "Alignment of multi-sensored data: Adjustment of sampling frequencies and time shifts," in *Proc. Comput. Cardiol. Conf. (CinC)*, Dec. 2019, pp. 2–5.
- [63] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [64] J. Koenig and J. F. Thayer, "Sex differences in healthy human heart rate variability: A meta-analysis," *Neurosci. Biobehav. Rev.*, vol. 64, pp. 288–310, May 2016.
- [65] M. Ferla and A. Graham, "Women slowly taking off: An investigation into female underrepresentation in commercial aviation," *Res. Transp. Bus. Manage.*, vol. 31, Jun. 2019, Art. no. 100378.
- [66] Ö. Yanikoğlu, S. Kiliç, and H. Küçükönel, "Gender in the cockpit: Challenges faced by female airline pilots," *J. Air Transp. Manage.*, vol. 86, Jul. 2020, Art. no. 101823.
- [67] H. A. Abhishekh, P. Nisarga, R. Kisan, A. Meghana, S. Chandran, T. Raju, and T. N. Sathyaprabha, "Influence of age and gender on autonomic regulation of heart," *J. Clin. Monitor. Comput.*, vol. 27, no. 3, pp. 259–264, Jun. 2013.
- [68] J. Zhang, "Effect of age and sex on heart rate variability in healthy subjects," *J. Manipulative Physiol. Therapeutics*, vol. 30, no. 5, pp. 374–379, Jun. 2007.
- [69] M. A. Almeida-Santos, J. A. Barreto-Filho, J. L. M. Oliveira, F. P. Reis, C. C. da Cunha Oliveira, and A. C. S. Sousa, "Aging, heart rate variability and patterns of autonomic regulation of the heart," *Arch. Gerontol. Geriatrics*, vol. 63, pp. 1–8, Mar. 2016.
- [70] B. Cinaz, R. La Marca, B. Arnrich, and G. Tröster, "Towards continuous monitoring of mental workload," in *Proc. Int. Workshop Ubiquitous Health Wellness*, Sep. 2010.
- [71] B. Cinaz, R. La Marca, B. Arnrich, and G. Tröster, "Monitoring of mental workload levels," in *Proc. IADIS Int. Conf. e-Health, EH, IADIS Multi Conf. Comput. Sci. Inf. Syst. (MCCSIS)*, 2010, pp. 189–193.
- [72] T. Heine, G. Lenis, P. Reichensperger, T. Beran, O. Doessel, and B. Deml, "Electrocardiographic features for the measurement of drivers' mental workload," *Appl. Ergonom.*, vol. 61, pp. 31–43, May 2017.
- [73] S. Miyake, S. Yamada, T. Shoji, Y. Takae, N. Kuge, and T. Yamamura, "Physiological responses to workload change. A test/retest examination," *Appl. Ergonom.*, vol. 40, no. 6, pp. 987–996, Nov. 2009.
- [74] R. Acharya, N. Kannathal, O. W. Sing, L. Y. Ping, and T. Chua, "Heart rate analysis in normal subjects of various age groups," *Biomed. Eng. OnLine*, vol. 3, no. 1, p. 24, Dec. 2004.



**ANDREA ALAIMO** received the degrees in aerospace engineering from the University of Palermo and the Ph.D. degree in aerospace structure from the University of Pisa. In 2003, he was a Visiting Student with the Queen Mary University of London. Between 2005 and 2009, he was a Postdoctoral Research Fellow with the University of Palermo. Since 2010, he has been with the University of Enna "Kore," at the beginning as an Assistant Professor. Since 2018, he has been as a Full Professor in aerospace structures. He is currently the Director of the M.A.R.T.A. Centre—Mediterranean Aeronautics Research and Training Academy.



**ANTONIO ESPOSITO** received the M.S. degree in aerospace engineering from the University of Palermo and the Ph.D. degree in aeronautical infrastructure technologies and management from the Kore University of Enna (UKE). He is currently an Assistant Professor in flight mechanics with the Faculty of Engineering and Architecture, UKE. He is also the Head of the Human Factor in Aeronautics Laboratory, M.A.R.T.A. Centre. The research areas currently concern the human factors in cockpit and cabin environments, with particular interests in warning systems for pilot's workload. The research topics in the flight mechanics field are focused on fault-tolerant control systems based on the configuration of simple adaptive control with parallel feed-forward compensator.



**PALMIRA FARACI** received the M.S. degree in psychology and the Ph.D. degree in public relations from the University of Palermo, Italy. She was a Visiting Scholar at Berkley University (California), attending the course measurement in education and social sciences. She is currently a Full Professor in psychometrics with the Faculty of Human and Social Sciences, Kore University of Enna, Italy. She is the author of several scientific publications in the field of psychometrics, collaborating with national and international research teams. Her research interests include psychological assessment and test construction, with a focus on factorial analysis, structural equation models, and item response theory.



**CALOGERO ORLANDO** received the M.S. degree in aerospace engineering from the University of Palermo and the Ph.D. degree in aerospace engineering from the University of Pisa. He is currently an Associate Professor in aerospace systems with the Kore University of Enna, Italy. He is also the Head of the Numerical Modeling and Analysis Laboratory, M.A.R.T.A. Centre. His research interests include the modeling and numerical analysis of aeronautical systems, the simple adaptive control SAC of aerospace systems, the stochastic robustness analysis, the PSO-based heuristic optimization procedures, and the subjective and objective evaluation of pilot workload and passenger comfort.



**GIUSY DANILA VALENTI** received the M.S. degree in psychology and the Ph.D. degree in social inclusion in multicultural contexts from the Kore University of Enna, Italy. She was a Visiting Graduate Researcher with the University of California at Los Angeles (UCLA), Los Angeles, from September 2018 to March 2019, attending the course multivariate analyses with latent variables and conducting research studies about acculturation among international students. She is currently a Research Fellow with the Faculty of Human and Social Sciences, Kore University of Enna, with a Research Grant for the Project RPASinAIR. Her research interests include personality assessment and psychological testing, with a focus on construction, validation, and adaptation of psychological measurement tools.

...