



A POSSIBILITY OF HUMAN VOICE ANALYSIS AS METHOD OF FUNCTIONAL MENTAL STATE EVALUTION IN AVIATORS

Oliver Dzvoník ¹

Ferdinand Tesař ², Jaromír Cmíral ² and Petra Tichá ²

¹ Institute of Aviation Medicine Prague

² Military Technical Institute Prague

Human voice

- Voice analysis is an every day thing and we all do it all the time.
- You now listen to my voice and you perform 3 different types of analysis, using your own human devices (your ears):
- You listen to the words I'm saying, trying to understand the textual message I'm delivering and its context.
- You listen to the uniqueness of my voice, the acoustical features that differentiate my voice from others. This is "Voice print".
- But you also do something else. You listen to me and you try to estimate: how stressed am I standing on this stage? how excited, how certain? What can my voice reveal about my personality?
- This is the field of "Emotion Detection".

Functional State of Operator (FSO)

FSO could be defined as:

- **multidimensional scheme of psychophysiological conditions, which allows**
- **human performance in relation to physiological and psychological demands.**



RTO NATO TECHNICAL REPORT (2004): Operator Functional State Assessment, AC/323(HFM-104)TP/48, Published February 2004)

Functional State of Operator - Definition

FSO results from:

- **operator`s characteristics**
(incl. his/her physical and mental fitness and health),
- **current work conditions**
and
- **interactions of operator**
with environmental
conditions.



Functional State of Operator – Basic state

- Basic state means **initial or average state of operator without workload, responsibilities and goals.**
- This state could be characterized as **type of state signature reflecting differences of psychological, physiological, cognitive and personality profiles.**



Functional State of Operator – Operational state

- Operational state reflects of operator`s functional state in time of task engagements in a specific work conditions.



Voice Analysis (VA)

- VA is a technology of recording and analysis of specific human voice characteristics.



Voice Analysis (VA)

VA technology is based on the principle, that

- **non-verbal and low-frequency contents of human voice, provides the information**
- **about physiological (health) and psychological (mental) state of speaking person.**



Some basic parameters of voice

- **Voice intensity (volume)**
- **Basic frequency**
- **Duration**
- **Spectrum of vocal tract**
- **Glottal resource** (direct relation to speaker)
- **Articulation profile of vocal tract** (direct relation to speaker)
- others, to be processed...



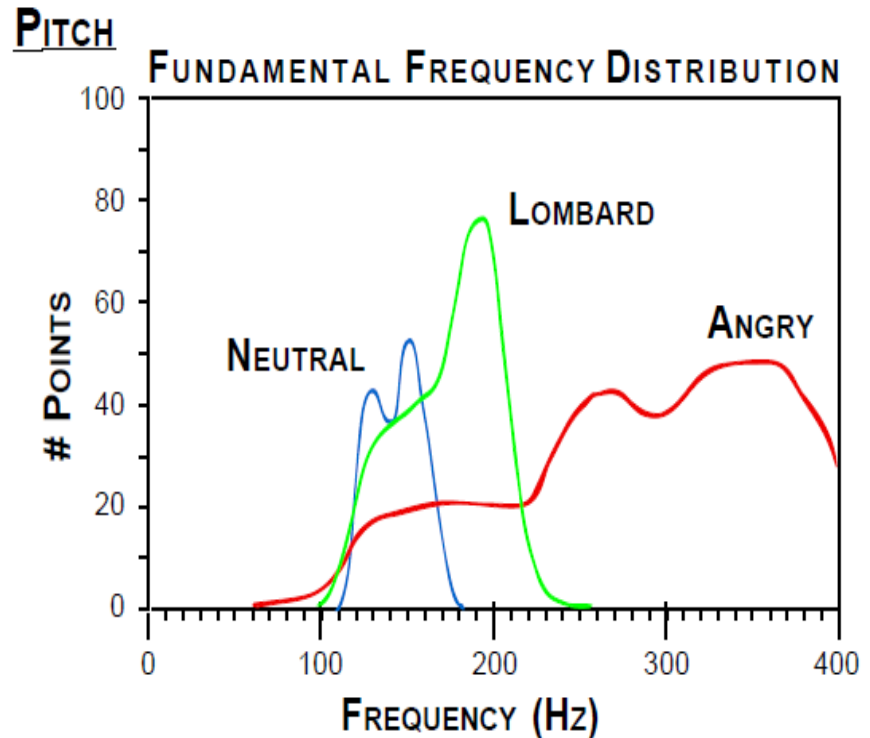
Basic parameters of vocal analysis

- Voice intensity –the average intensity is measured when environmental noise increases (Lombard effect), anger or some type of higher loads.
- Basic frequency - most acknowledged stress assessment parameter, with an assessment of their profile, distribution and dispersion.
- Duration – average duration of word is significant individual indicator of speech in different conditions (e.g. speed, clarity, anger, Lombard effect and noise, etc.)
- Spectrum of vocal tract – the vocal maximum area and voice spectrum shows significant changes in different types of stress conditions.



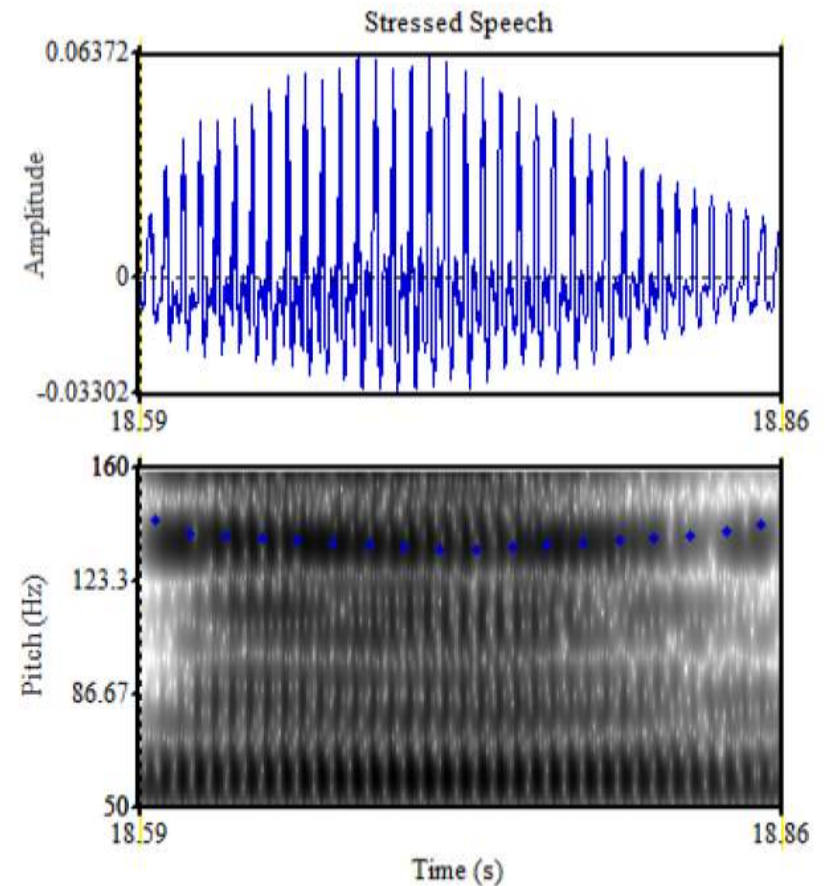
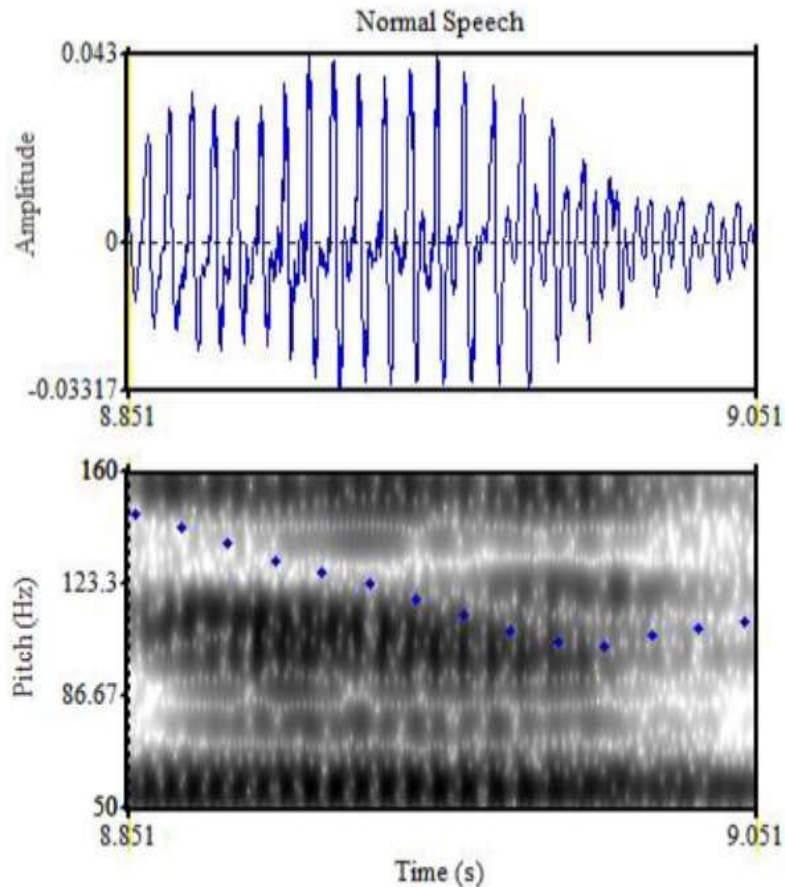
Basic frequency distribution

- Basic frequency by neutral speech, and reactions of speech on Lombard effect a anger as a state.



	<u>NEUTRAL</u>	<u>LOMBARD</u>	<u>ANGRY</u>
MEAN (Hz)	145	160	253
STAND. DEV.	15	24	95

Spectrograms of normal human speech and human speech under stress (Sondhi et.al., 2015)



Stress indicators in human voice

Sondhi et al. (2015) carried out a study on stress and speech changes and indicators:

- in case of emotional stress exposition there are **clear changes or indicators in human voice**,
- the average values of voice **frequency of formants (F1-F4)** were derived as **significant indicators of psychological state - emotional stress**.

	Normal speech	Stressed speech
Mean Pitch (Hz)	185.38	216.90
Formant F1 (Hz)	335.93	269.03
Formant F2 (Hz)	2256.22	1941.67
Formant F3 (Hz)	2860.49	3142.49
Formant F4 (Hz)	4267.45	3987.12
Jitter (%)	1.36	1.18
Shimmer (%)	8.60	6.82

Human voice analysis and fatigue detection

- Several speech manifestations and voice changes under fatigue were described and focused on discrete voice characteristics (encompassed e.g. word repetitions, oscillation and duration). (Vollrath, 1994).
- Spectral parameter changes in speech were associated with effect of alcohol use (Brenner & Cash, 1991) and hypoxia (Saito et al., 1980).
- Significant effects of circadian biases on human voice characteristics were observed in many other studies (Roth et al., 1989; Whitmore & Fisher, 1996).

Personality and human voice (speech)

A verbal expression is widely perceived as one of the essential attributes of human behavior (Crow, 2000; Premack, 2004).

A relations between speech (voice) and individual personality characteristics were examined in many studies (e.g. Chung & Pennebaker, 2008; Tausczik & Pennebaker, 2010),

But relations between communication patterns and personality dimensions remains still a challenge for the next research.



Mental health and human voice (speech)

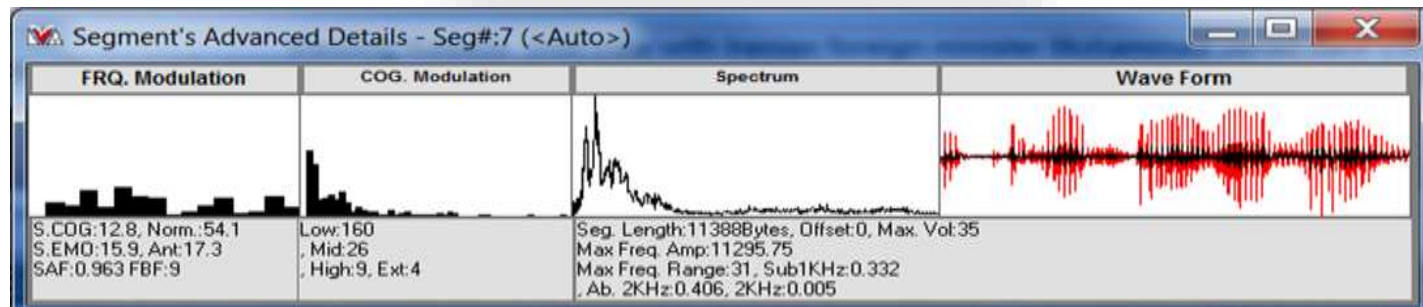
A possibility of mental diseases diagnostics through perception or acoustic voice analysis would be welcome.



- Moore et al (2003) evaluated **patients suffering of clinical depression based on speech features extracted from glottal waveforms.**
- Alpert et al (2001) found out that clinical **depression was essentially related to acoustic parameters e.g. frequency and prosody** in depressive patients.
- Mund (2006) – carried out a study on **acoustic voice patterns of depressive patients and process development from early stages till the treatment** on the scale of extracted and manifested symptoms.
- These findings also **suggest to apply an acoustic measurements of patient voice for the objective depression assessment** purpose.

Layered Voice Analysis (LVA)

- **Layered Voice Analysis™ – LVA™** by Nemesysco is designed to detect and measure a wide range of emotions and cognitive states, collect them over time period and estimate the tested parties emotional profiles based on the aggregated data.



LVA – measured emotional/cognitive categories

1. **Emotional:** measures excitement level. Normal values are around the middle of scale.
2. **Logical:** measures cognitive activity level. Low values indicate lower conflict level (higher certainty) and lower processing effort. High levels indicate conflict and uncertainty. Normal values are around the middle of the scale.
3. **Hesitant:** measures the level of “openness” and self-criticism. Normal values are around the middle.
4. **Stressed:** Level of „fear“and self protecting state
Normal values are around the middle-low part of the scale.
5. **Energetic:** measures level of engagements, based on the Energy value. Low values indicate tiredness (fatigue) and lack of desire to engage. High levels may anger or joy but in case, the increase of energy typically shows an increased level of engagements.
6. **Thoughtful:** measures level of mental effort. Normal values are around the middle-lowpart of the scale.



LVA

This window presents several graphical elements presenting different aspects of the analysis results in real-time.

Online Test

STOP File

Batch Process

1.wav

Real-Time View

Call in-depth view

Call Logger/ Database

Network Settings

License Manager

Close

nemesysco | QA5
Call Center Quality Assurance

Emotion Call Priority: 76

Define Real-Time alarm settings:

☒ Play segments as they are created

☐ Automatically start playing in High Emotion-Priority

☐ Silent analysis

☐ Use RIGHT channel in STEREO files


ANGER 20

Yellow = Stress, Red = Anger


Green = Content, Red = Upset, White Line = Call Priority


Green = Overall mental activity, Red = Emo/Cog, Yellow Line = Energy


Nemesysco's Emotional Diamond™




Emotional Profile:


Emotional: 

Logical: 

Hesitant: 

Stressed: 

Energetic: 

Thoughtful: 

Measured variables and parameters by vocal analysis

Variable	Description	Parameter
SPT	Emotional level	Average number of „thorns“
SPJ	Cognitive level	Average number of „plateaus“
JQ	Stress level	Standard error of plateau length
AVJ	Level of thinking	Average length of plateau
SOS	„Say or stop“ fear or reluctance level	?
FJQ	Imagination	Uniformity or low-frequency
FMAIN	Concentration level	Most important frequency range
FX	Concentration level	Frequency exceeded FMAIN
FQ	Insincerity	Uniformity frequency spectrum
FLIC	Hesitation and conflict thoughts	Frequency of harmonic spectrum
ANTIC	Anticipation	?
SUBCOG	Unconscious cognitions	?
SUBEMO	Unconscious cognitions	?

Our research attempt

Subjects: A small group of **6 fighter pilots** were repeatedly tested **during several simulated tactical exercises.**

Methods: A **voice analysis from communication, heart rate (HR) and personality survey (ALAPS)** were simultaneously evaluated and correlated.

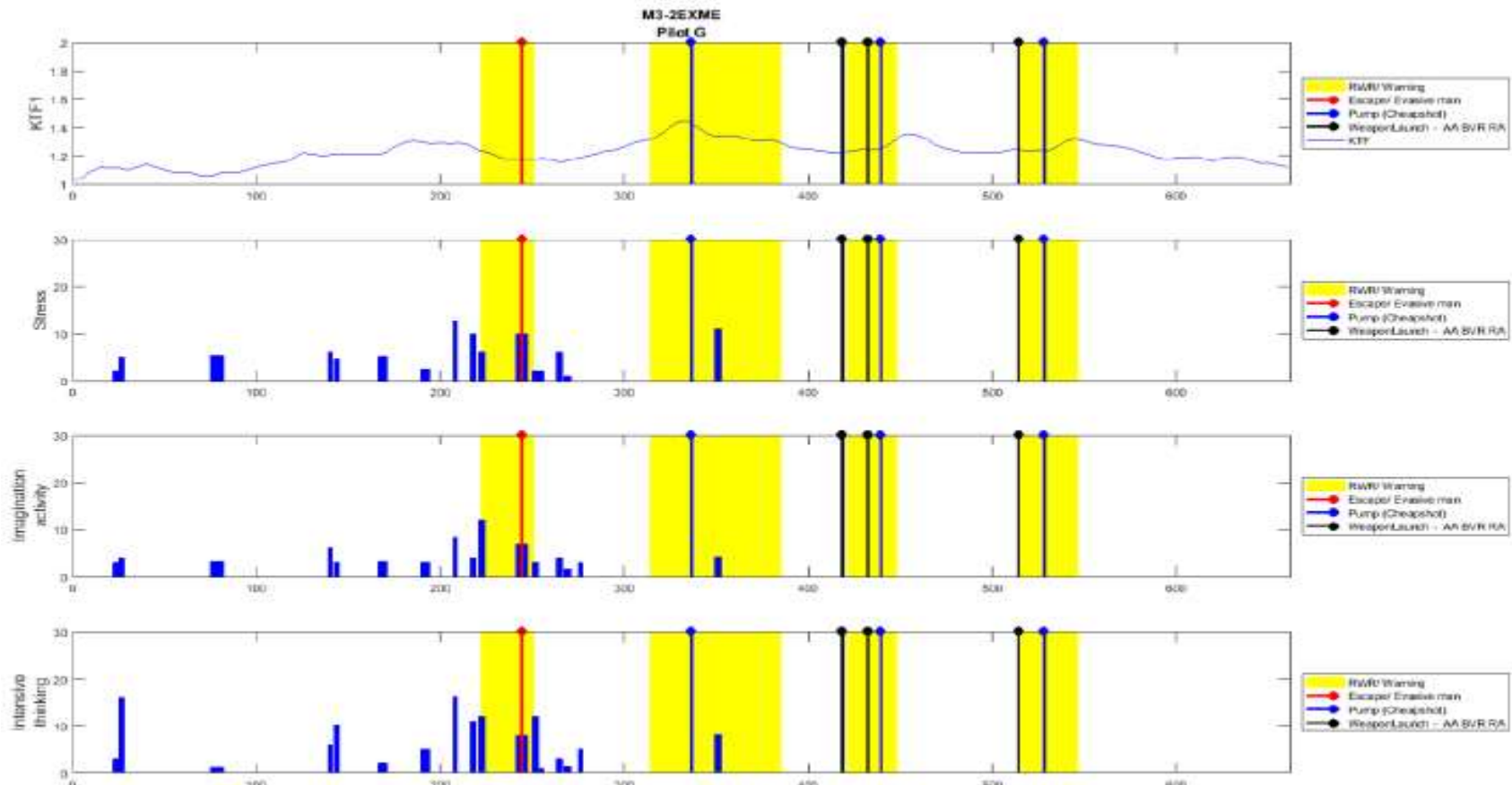
An additional approach evaluation was based on common context of eye tracking and EEG measurements (not yet included).

Our research attempt

Design:

- As the programme cannot make the analysis of the short records (pilot's communication is very short - several seconds only) two methods have been used.
- In the **first the individual records have been repeated to one longer record** (the voice analysis for each speech).
- **In the second method the individual records have been cumulated to one record** (one voice analysis for the whole exercising).
- The results obtained by the first method didn't correspond with the other physiologic values and so this method has not been used for the evaluation.
- **The second method seems as the applicable method for the evaluation.**

Comparison of HR with voice parameters (stress, imagination activity, intensive thinking – obtained by repeated records) in fighter pilot during simulated mission



Legend:

Yellow – RWR Warning – Aiming by enemy aircraft

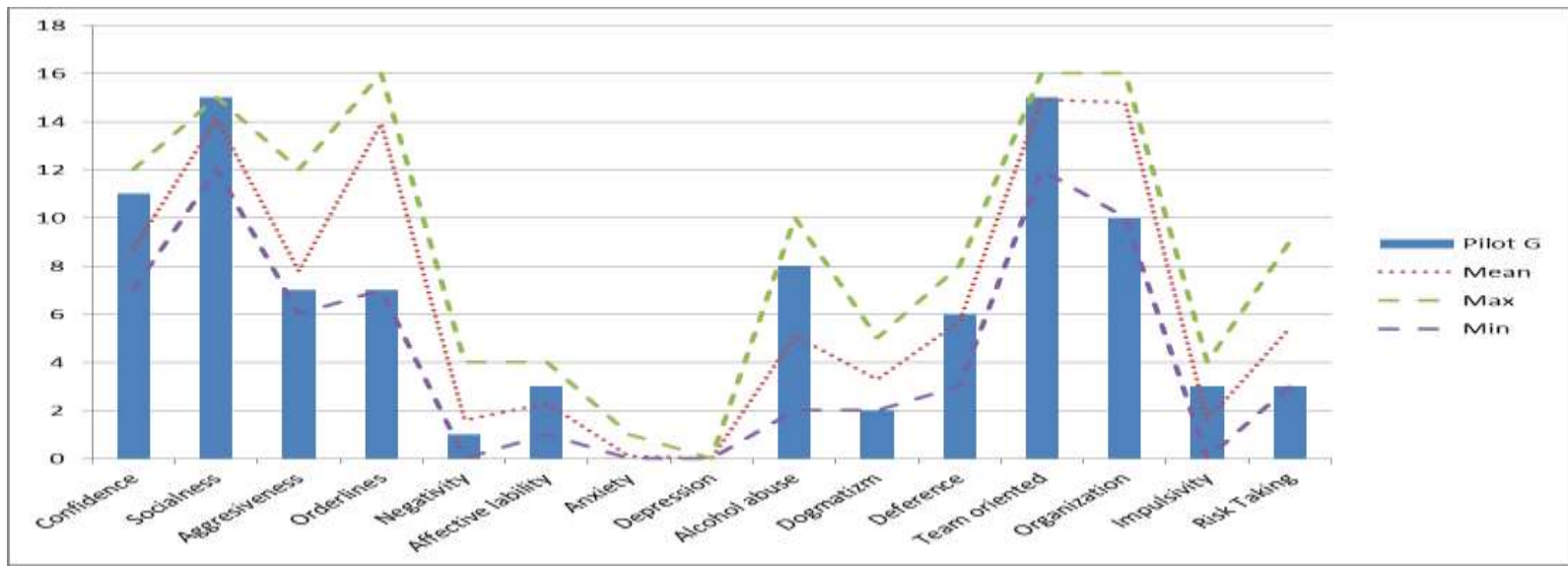
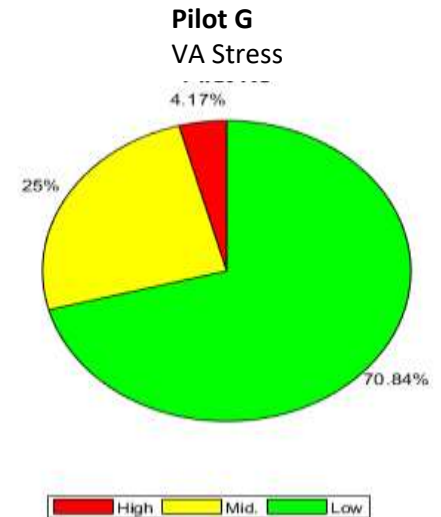
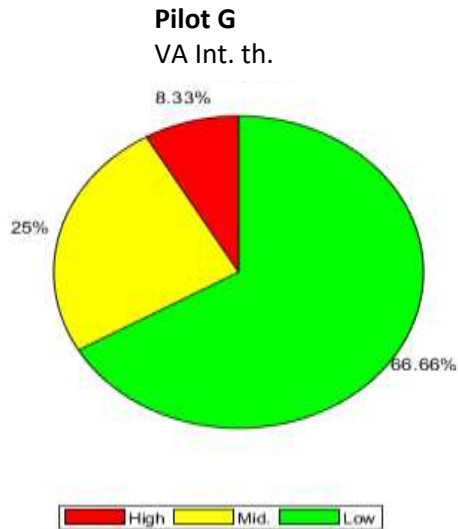
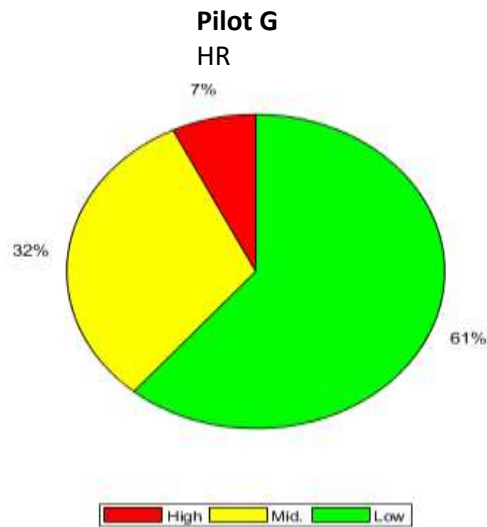
Red – Escape/Evasive Maneuver

Blue – Pump (Cheapshot) – a pilot finished lead missile on target (missile did not yet come to self leading)

Black – Weapon Launch – rocket launched

Blue Curve – KTF – HR frequency coefficient as multiplication of basic/quiet HR before measurement

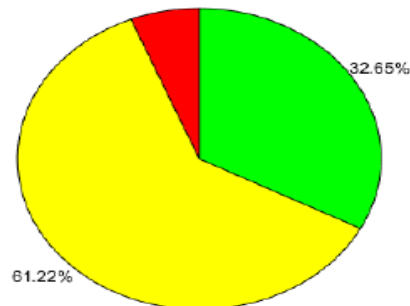
Voice analysis (VA) - comparison of cumulated „HR-stress“ , „VA Stress“, and „ VA Intensive thinking“ (identical situation)



Voice analysis (VA) - comparison of cumulated „Stress“ among different pilots during missions

Commander (H)

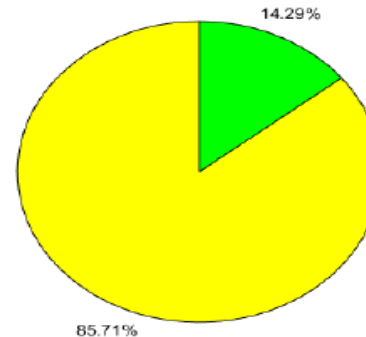
M3-3EXME
Stress
Pilot VB1 (H)
6.12%



High Mid. Low

Deputy comm. (E)

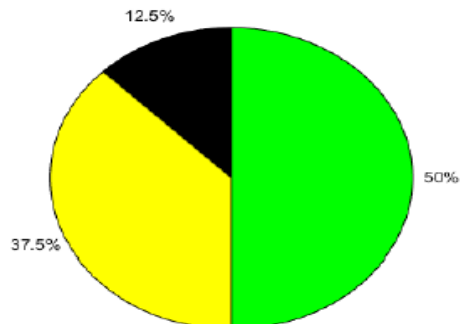
M3-3EXME
Stress
Pilot B2 (E)
14.29%



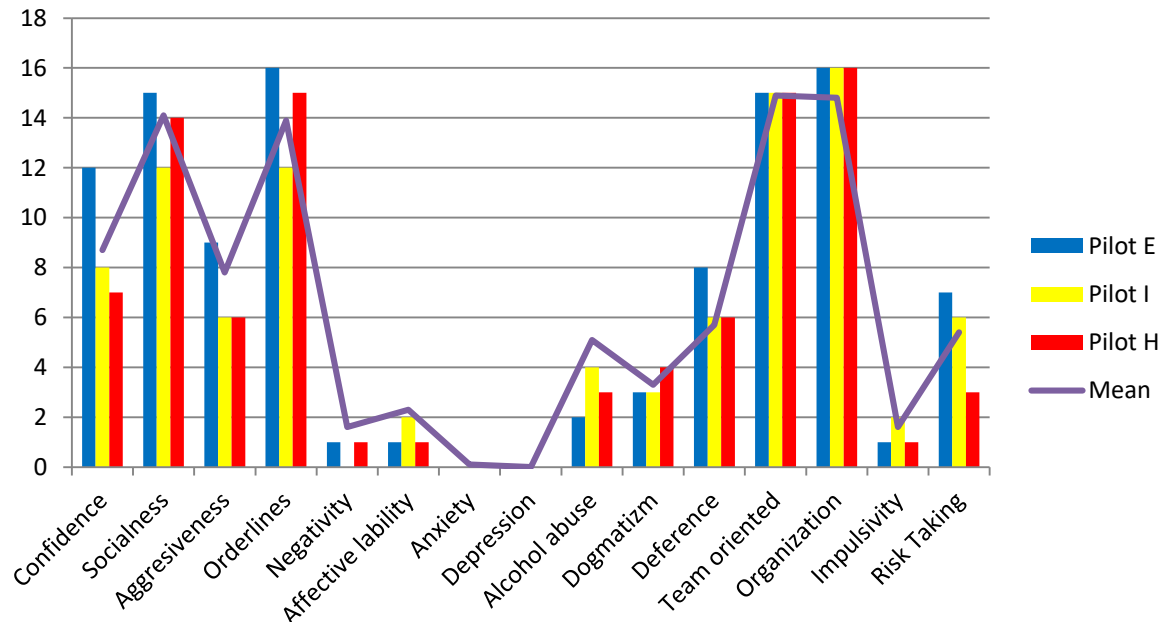
Mid. Low

Pilot (I)

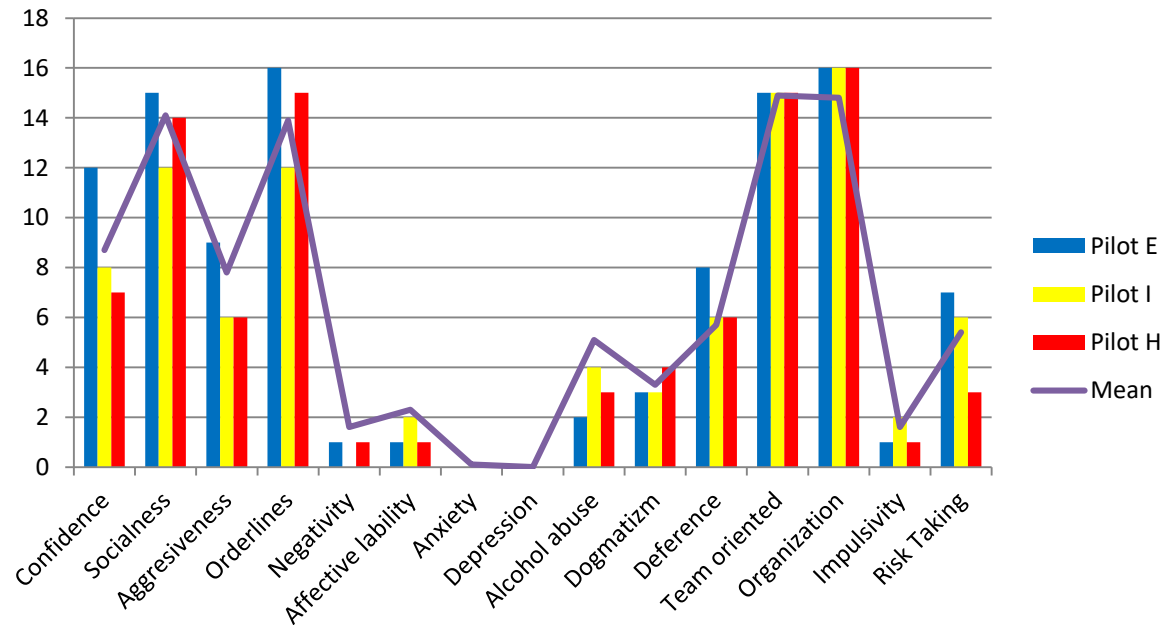
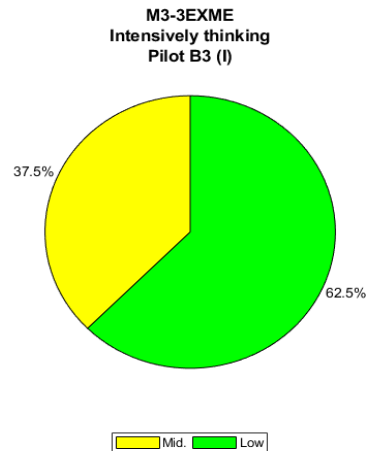
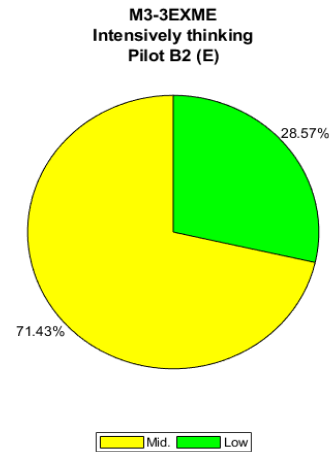
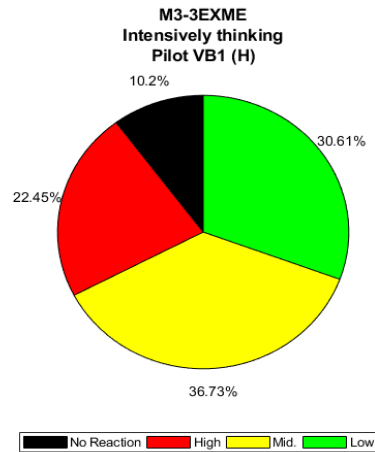
M3-3EXME
Stress
Pilot B3 (I)
12.5%



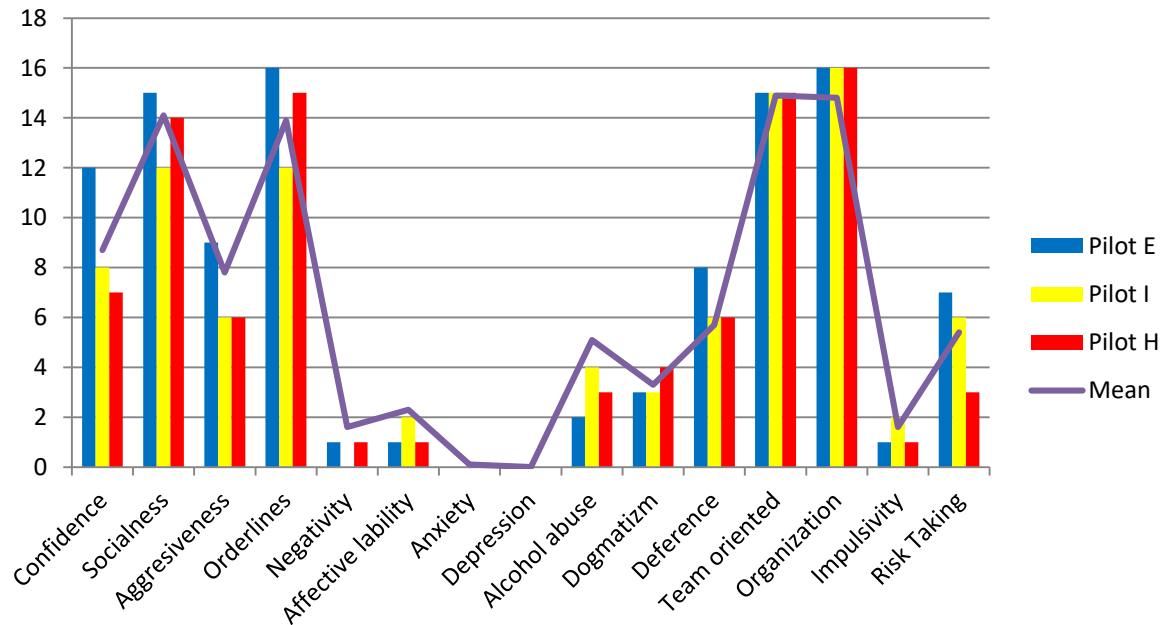
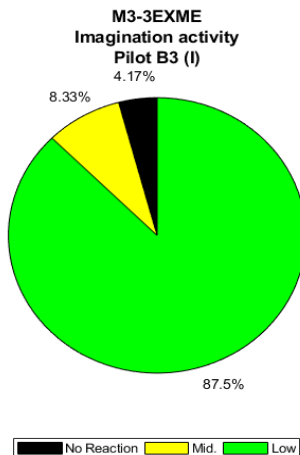
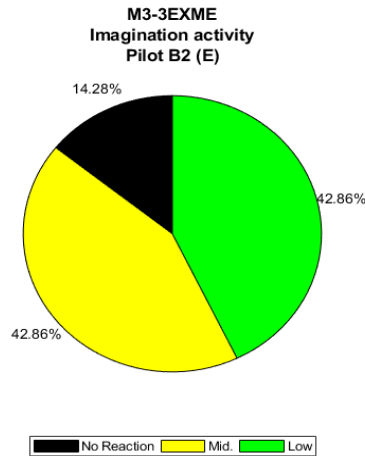
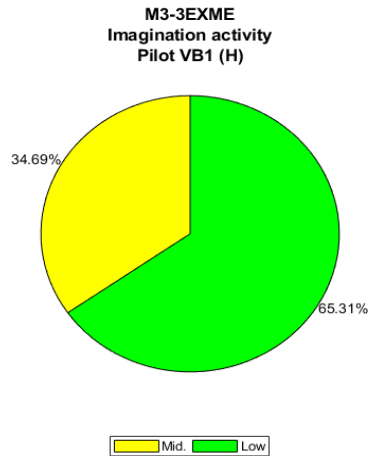
No Reaction Mid. Low



Voice analysis (VA) - comparison of cumulated „VA Intensive thinking“ among different pilots during missions



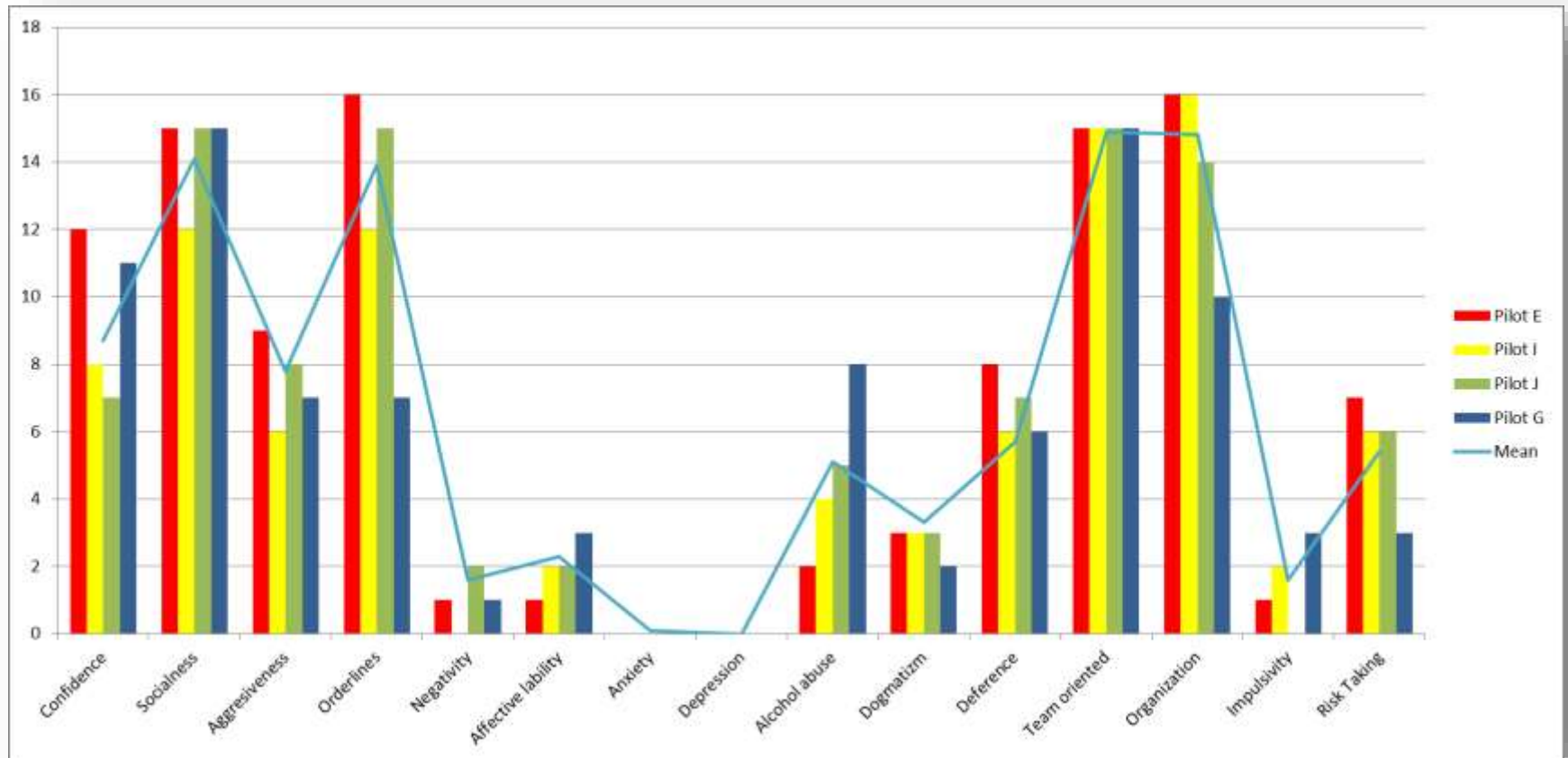
Voice analysis (VA) - comparison of cumulated „VA Imagination Activity“ among pilots during missions



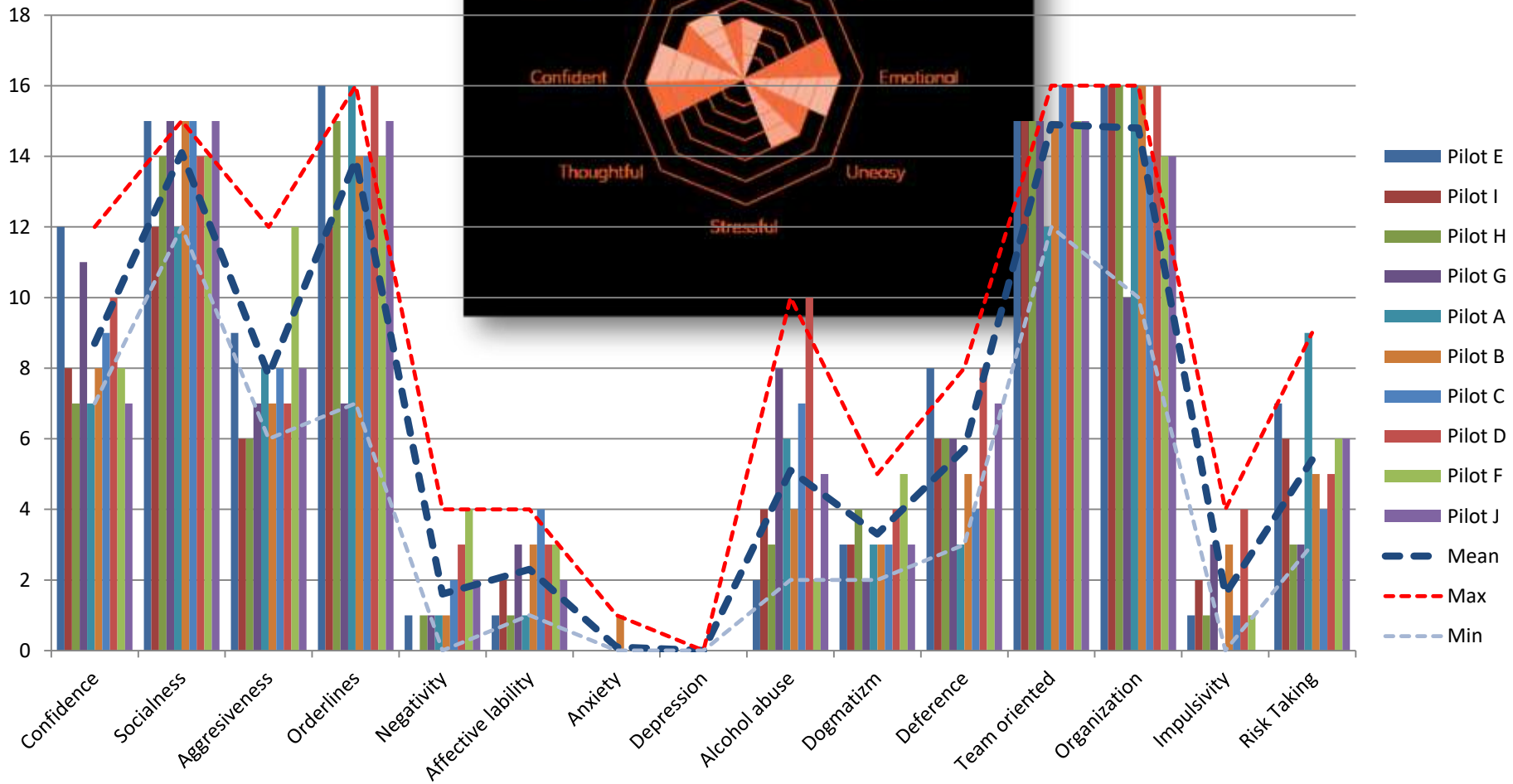
LVA comparison of stress, intensive thinking and imagination activity levels among pilots during the mission

pilot	Stress					Imagination activity					Intensively thinking				
	average (marks)	without (%)	high(%)	middle (%)	low (%)	average (marks)	without (%)	high(%)	middle (%)	low (%)	average (marks)	without (%)	high(%)	middle (%)	low (%)
J	14,31	0	9,52	83,33	7,14	7,1	2,38	0	14,29	83,33	10,5	0	2,38	59,52	38,1
E	18,5	0	44,44	55,56	0	8,39	0	5,56	16,67	77,78	21,17	0	66,67	33,33	0
G	4,52	24	0	20	56	3,68	32	0	8	60	5,52	24	0	32	44
I	18	0	28,57	71,43	0	4,86	14,29	0	0	85,71	12,71	0	14,29	42,86	42,86
D	10,11	0	2,22	57,78	40	7,13	4,44	0	22,22	73,33	8,96	2,22	4,44	37,78	55,56
B	5,5	0	0	0	100	3,7	30	0	0	70	2,8	50	0	0	50

ALAPS – profiles comparison of pilots E,I,J,G



Our study on voice analysis validity to be continued...



Conclusion I

- It seems to be more likely that **relations among several psycho-physiological parameters and operational load is not always linear and/or immediate.**
- It could be also affected/influenced by **anticipation or post-operational stress responses or both.**
- Some parameters are ***relatively sensitive*** (e.g. HR, VA stress, VA intensive thinking) ***towards sudden and/or rapid changes*** and some of them are not.

Conclusion II

- **Human voice analysis seems to be promising psycho-physiological method** in measurement of cognitive and emotional workload or stress in aviation operator's professions (e.g. pilots and air traffic controllers).
- **Workload, stress and fatigue are worth to be explored** (not only from their relations on human performance) but also **from the view of their effects on aviation mental health and fitness.**

Conclusion III

- If the next research we shall be verifying a closer correlation between voice analysis, stress, fatigue and others psychological states.
- The **voice analysis could became a new non-invasive and very comfortable method for functional mental state testing.**
- This method and approach don't require any special sensors or additional equipment and **aviation operators would be evaluated on-line or in real time.**

Thank you for your attention!



Dr. Oliver Dzvonik, PhD. (Lt. Col. Ret.),
Aviation and Clinical psychologist
Institute of Aviation Medicine
Department of Psychology and Psychiatry
Gen. Píky 1, 160 60 Prague, Czech Republic

dzvonik@ulz.cz

www.ulz.cz

Literature

RTO NATO TECHNICAL REPORT (2004): Operator Functional State Assessment, AC/323(HFM-104)TP/48, Published February 2004, ISBN 92-837-1111-4 <https://www.cso.nato.int/Pubs/rdp.asp?RDP=RTO-TR-HFM-104>

Sondhi S., Khan M., Vijay R., Salhan A. K.(2015) Vocal Indicators of Emotional Stress, International Journal of Computer Applications (0975 – 8887) Volume 122 – No.15, July 2015
<http://research.ijcaonline.org/volume122/number15/pxc3905056.pdf>

Luig J., Sontacchi A. (2012): A speech database for stress monitoring in the cockpit, Journal of Aerospace Engineering, Proc IMechE Part G:J Aerospace Engineering 0(0) 1–13! IMechE 2012, DOI: 10.1177/0954410012467944
http://www.utdallas.edu/~hynek/citing_papers/Luig_A%20speech%20database%20for%20stress%20monitoring%20in%20the%20cockpit.pdf

Liberman, A. (2003) Apparatus and Methods for Detecting Emotions. US Patent 6638217 B1. <http://www.freepatentsonline.com/6638217.html>

Keller B. (2005). Speech prosody, voice quality and personality. Logopedics Phoniatrics Vocology, 5, 72–78. doi: 10.1080/14015430500256543

Polzehl T. Moller S. Metze F. (2015): Automatically Assessing Personality from Speech retrieved from: http://www.cs.cmu.edu/~fmetze/interACT/Publications_files/publications/icsc_FM.pdf

Steeneken H.M.J., Hansen J.H.L. (1999): SPEECH UNDER STRESS CONDITIONS: OVERVIEW OF THE EFFECT ON SPEECH PRODUCTION AND ON SYSTEM PERFORMANCE,
<https://ieeexplore.ieee.org/iel4/6110/16342/00758342.pdf>