

SUPPLEMENTARY MATERIALS

TABLE I: Summary physiological measures contributing to team performance analysis

Individual level	Processed	Team-level	team-size	Performance
EEG				
<i>alpha</i> , <i>beta</i> , <i>gamma</i> , <i>theta</i> and <i>delta</i> frequency spectrum	integer levels indicating $EEG_E$ and $EEG_{WL}$ all team members were combined to form patterns	Sliding window based Entropy was estimated from temporal $EEG_E$ patterns	six	Discriminant function analysis revealed more experienced teams had higher mean entropy compared to the less experienced teams[52]
	$EEG_E$	1-second lag transition matrix to detect attractor basins from $EEG_E$	six	Attractor basins of experienced teams were different from inexperienced teams [51]
	$EEG_{WL}$	From temporal $EEG_E$ , sliding window based Entropy was constructed	six	Inexperienced teams were characterized by more irregularities in the Entropy compared to experienced teams [51]
		From temporal $EEG_{WL}$ , sliding window based Entropy was constructed	dyadic	mean of the Entropy correlated with mistakes [54]
		averaged interconnectivity map, strength density (ICD), clustering coefficient (IALD)	dyadic	Interconnectivity map showed strong connections in $\alpha$ and $\theta$ bands during harder phase of task which also correlated with reduced performance, and contrasting results were observed in easy conditions. This phenomenon was also complemented by higher values of ICD and IALD in specific region of brain [55]
		ICD, IALD	dyadic	ICD of $\alpha$ and $\theta$ bands was significantly higher in real teams during high workload phase of the task compared to pseudo generated teams and IALD showed the region of brain activity [27]
FNIRS	oxy-hb, deoxy-hb	Pearson CC used to establish inter-brain synchronicity (IBS) between participants brain regions	dyadic	IBS decreased post negative feedback[47]
	oxy-hb	wavelet coherence used to calculate average IBS	dyadic	change in coherence correlated with cooperation phase of the task [39], and was significant in mixed gender groups
ECG				
IBI		WC (High frequency and low frequency components)[24]	dyadic	
		cross-correlation, cross-fuzzy entropy (CFE)	dyadic	High score correlated with low cross-correlation between team members IBI, and CFE negatively correlated with coupling [53]
	mean-IBI and from HRV MSD, RSA and log-RSA	SM, DA, IDM and CC	4	correlation was significantly high for high performing teams compared to low performing teams [17]. Trend-level significant difference was observed with DA
ECG and ICG combined analysis	log-RSA association with PEP and LVET, respectively to obtain a composite score	cross-correlation to derive team level index	dyadic	From ordinary least squares, composite scores demonstrated significant associations to team error as well as distinctly associated with individual operators [65]
EDA				
EDA		$R^2$ , lyapunov exponents, synchronisation weight	dyadic	Pearson-r showed correlation between synchronisation weight and false alarm rates [64]
	SCR	DA	dyadic	DA was related with higher performance levels [24]
	SCR	SM	dyadic	SM was related performance in resource management [24]
	SCR	WC (High frequency and low frequency components)[24]	dyadic	
	SCR	MWE	dyadic	MWE significantly distinguished fail and pass conditions [26]
				higher SCR levels shared relation with failed condition visually distinguishing from a passed condition [26]
	SCL and SCR			SCL and NSCR levels tended to be higher in positive outcome groups [104]

Individual level	Processed	Team-level	team-size	Performance
EMG				
OO, ZM			dyadic	Have been employed to observe participants responses post task outcome [40]
SM,CS			dyadic	To observe emotional changes due to chance to employ performance enhancing routine [50]
Eye-tracking				
Eye-gaze	fixations	CRQA, %RR	dyadic	%RR decreased as team performance increased [67]
Communication				
speech transcript	LSA	semantic vector length	six	
	LSA cosine	hierarchical clustering	six	More experienced teams clustered from less experience teams [52]
	Non-content based	CCR index	three	Experienced teams showed significantly higher CCR index compared to the inexperienced teams [69]
Actions				
accelerometer voltage recordings		MdRQA %Determinism and %Laminarity	dyadic	During the egalitarian condition, lower values corresponded with better looking designs [48]
video		de-synchronisation	dyadic	de-synchronisations were less among elite teams who also took shorter duration to complete the task [45]

TABLE II: Summary multimodal measures contributing to team performance analysis

Individual level	Processed	Team-level	team-size	Performance
	Interaction: idle, construction and execution; Speech: silence, controller speaking and contributor speaking; body movement: low and moderate	Single mode: Only interaction measures, dual mode: All combinations of Interaction and Speech measurements, multimode: Combination of measurements from all three individual modes	3	Only specific combination of measurements demonstrated significant correlation with task scores [68]
EDA, ECG, Respiration, EMG, EEG			2	Cooperation and collaboration [94]
EDA, ECG, Respiration, EMG			2	Cooperation versus competition [33]
Facial expression, ECG, Linguistic Inquiry	Facial expression: joy, anger, fear, positivity and confusion; ECG: HRV, Speech: word count, sadness, general, insights and cognitive processing	Facial expression and Speech: mean counts and HRV: RMSSD	2	Impact of emotions[1]
Speech features and ECG	Speech: fundamental frequency and peak slope, ECG: cardiac output and total peripheral resistance		2	Only relation to stress, and no relation to team performance [25]
audio, video, ECG, EDA			2	Annotated multimodal data for analysing collaboration versus affective states[105]
Speech, body movement and EDA	Signals from individual mode was z-score normalized averaged across participants.	MdRQA was used to estimate the %recurrence	2	No significance between recurrence and task score [38]
EEG, Eye-tracking			2	No collective analysis observed. Expert teams were able to quickly identify areas of interest compared to novice teams by approximately 3 times longer [36]
ECG, EEG, Respiration and GSR	features were extracted for each modality using statistical and signal processing approach		2	Classifiers trained with bench-mark data were used to classify each participants modality features. Multimodal and individual classification was performed to observe the cognitive state of participant[29]. Team of 2 members participated together in flight simulation task, however, there was no team level analysis.
gaze and speech	gaze measures were used to compute cognitive load and windowed-entropy, and speech audio: four different prosody tones	proportional similarity between gaze-entropy and cognitive load of the participants	2	LSTM to predict learning gain and post-test score. Input data comprised of gaze and audio provided best prediction of post-test score [30]
physical position, speech and interaction, EDA, actions	For every participant, 5 location data, binary coded presence or absence of speech and listening, respectively and EDA peaks was acquired	Each modality form every participant, was compiled into multimodal matrix	2-3	task specific analysis which majority depended on the patient's needs [37]