

Objective Characterization of Human Behavioural Characteristics for QoE Assessment: A Pilot Study on the Use of Electroencephalography Features

Khalil ur Rehman Laghari*, Rishabh Gupta*, Jan-Niklas Antons†, Robert Schleicher†, Sebastian Möller† and Tiago H. Falk*

*INRS-EMT, University of Quebec, Canada †Quality and Usability Lab, Berlin Institute of Technology, Germany

Abstract—Quality of Experience (QoE) is a human-centric paradigm which produces the blue print of human behavioral states such as perception, emotion, cognition and expectation. Recent advances in neurophysiological monitoring tools have facilitated the study of frequency, time and location of neuronal activity to an unprecedented degree, as well as opened doors to a better understanding of human cognition, emotions and overall behavioral systems. These neurophysiological insights may provide more accurate and objective characterization of QoE metrics. This paper seeks to investigate neuronal activity due to different stimulus quality levels in order to understand human behavioral states at the neural level. An electroencephalography (EEG) feature was computed based on the coupling between so-called delta and beta EEG frequency bands, which has previously been linked with negative behavioral characteristics (anxiety, frustration, dissatisfaction). The result indicates an increase in delta and beta coupling with a decrease in the speech quality levels. Additionally, neural correlates of a subjective affective scores (arousal and valence) were also computed and shown to be inversely proportional with EEG feature. These preliminary findings corroborate that emotions play a significant role in human quality and QoE perception.

I. INTRODUCTION

Quality-of-Experience (QoE) is a fast emerging multi-disciplinary field which provides an assessment of human emotions, perceptions, cognition and behavior with respect to a particular product, service or application [1]. QoE is a promising solution to multimedia service providers and vendors who, in this very competitive environment have been yearning for rich customer-centric solutions. QoE measurement, however, is challenging due to the variability and complexity of human behavior, as not all humans have similar preferences, feelings or perceptions about a particular service or product. Furthermore, user perceptions and preferences continuously change over time. The challenge is how to better understand human behavioral states and transform them into meaningful data.

Human behavior is shaped by internal and external factors. Internal aspects, for example, include biological, psychological, and cognitive factors, while external aspects are related to social, economic, and technical factors [2]. In psychology, incentive theory describes external influences (e.g., environment, and rewards) as motivating factors for humans to do things, while on the other hand drive theory holds on that it is in fact the inner human biological need which is the driving

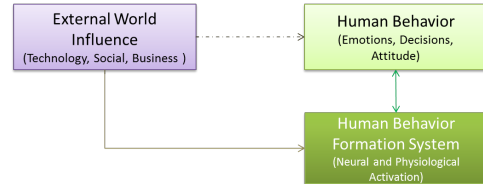


Fig. 1: Conceptual view of the human behavioural formation system

force behind human actions. Both internal human factors and external factors influence human behavior as depicted by Figure 1. External world elements comprising technology, business, and contextual aspects may exert some influence, triggering the human neuro-physiological system, which in turn will develop behavioral traits linked to QoE.

Conventionally, subjective methodologies proposed by the International Telecommunication Union (ITU-T) have been employed for capturing Quality of Experience factors such as user perception, emotion and frustration. Gradually, however the QoE research community has been moving towards exploring neuro-physiological tools to gain a better understanding of human neuronal activities and physiological responses responsible for human behavior formation. Neurophysiological insights may be obtained via neuroimaging techniques, such as electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and near-infrared spectroscopy (NIRS), as well as body area sensors and networks. EEG/MEG relies on measuring the electrical/magnetic activity in the brain. Whereas fMRI and NIRS are based on tracking blood flow (and correlates) that accompanies neuronal activity. While EEG and MEG provide data with high time resolution (in the order of milliseconds), they provide limited spatial resolution. In contrast, fMRI and NIRS provide good spatial resolution but relatively poor temporal resolution [3], [4]. For review on these neuro-imaging characteristics, readers can refer to [5].

Electroencephalography (EEG) features have shown to provide useful insights for QoE characterization. For example, the $P300$ event-related potential (ERP) signal, which occurs $300ms$ post stimulus presentation, has shown to be a useful EEG feature in characterizing the quality of text-to-speech

(TTS) [6], video [7], and audio-visual systems [8]. Another EEG feature based on wavelet theory was shown to be useful for detecting artefacts in image and video systems [9]. Another common EEG feature is EEG sub-band spectral power, which was recently shown to be useful in characterizing emotional states of users watching music clips [10]. On the other hand, the clinical research literature has suggested to use cross-frequency coupling as a mean of characterizing human emotions, mental activation status, and cognition. For example, in [11], the delta ($1 \leq \delta \leq 4$) and the beta ($13 \leq \beta \leq 30$) frequency sub-bands have been linked to behavioral inhibition states (anxiety and frustration). As such, this paper will explore the use of $\delta - \beta$ coupling as a correlate of speech QoE perception.

The remainder of this paper is organized as follows, Section II presents background and motivation of this pilot study, Section III describes the experimental methodology, and in Section IV, subjective and neurophysiological results are presented. Lastly, Section V presents the discussion and concludes the paper.

II. BACKGROUND AND MOTIVATION

ITU-T has proposed several subjective methods for measuring the perceptual Quality of Experience of audio [12] and video services [13]. On the other hand, the self assessment manikin (SAM) has been normally used to measure human affective states [14]. SAM is a non-verbal pictorial assessment technique that directly measures the valence, arousal, and dominance associated with a person's affective reaction to a wide variety of stimuli [14], [15]. SAM scales are depicted by Figure 2 for arousal and valence categories. For valence, SAMs range from a smiling, happy Manikin to a frowning, unhappy Manikin; for arousal, SAMs range from very excited with eyes-open to a sleepy with eyes closed Manikin [16]. Many empirical works have repeatedly confirmed that valence, and arousal are pervasive in organizing human judgments for a wide range of perceptual and symbolic stimuli, for instance, audio [17] and video [18]. As such, in this study focus will be placed on these two dimensions (arousal and valence).

Human behavioral characteristics can be broadly classified into two major categories: Behavior inhibition system (BIS) and behavior activation system (BAS). Gray, in his seminal work [19], suggested that the human behavioral inhibition system and the human activation system are reciprocally related to each other. BIS regulates aversive motivational systems, and promotes avoidance behavior and feelings of anxiety, frustration, uncomfot and dissatisfaction, while the BAS regulates appetitive motivation, such as reward, pleasure, and satisfaction. To understand the neural basis of these behavioral states, normally neuroscientists investigate the prefrontal cortex (PFC), and temporal cortex as they are known to be related with affective processing. In [20], it is suggested that the temporal cortex might also be important for the assignment of percept and affective meaning, and thus might perform a key role in experience of emotions, specially, anterior temporal lobe has been shown to perform a principal role in eliciting a subjective emotional experience such as valence [21], [22].

On the other hand, the prefrontal cortex generates the affective facial expressions and the control of affective behavior, thus it represents the expression of emotion [20]. The medial frontal cortex (MFC) is an important part of PFC and it has been linked with arousal [23].

Normally, EEG frequency sub-band powers are computed to assess BIS and BAS characteristics. It has been found in several studies that $\delta - \beta$ coupling is related to behavioral inhibition system characteristics. In [11], authors have found higher positive correlation between delta and beta powers in subjects with higher baseline level of salivary cortisol, (the steroid hormone directly associated with anxiety). In [24], authors showed that the correlation between midfrontal delta and beta spectral power increased in healthy male subjects with an increase in anxiety and behavioral inhibition. It has also been hypothesized that $\delta - \beta$ coupling reflects higher cortical arousal in frustrating situations [25]. And they also found that $\delta - \beta$ coupling is very sensitive to external influences since it allowed discriminating between good and bad performance condition. Motivated by these promising insights, goal of this paper is to investigate if $\delta - \beta$ coupling can be used for QoE characterization, as a correlate of affective user state (valence, arousal).

III. EXPERIMENTAL METHODOLOGY

A. Participants

Fifteen subjects participated in this study (eight female, seven male; mean age = 23.27 years; SD = 3.57; range = 18 – 30); all of them were fluent English speakers. All participants reported normal auditory acuity and no medical problems. Participants gave informed consent and received monetary compensation for their participation. The study protocol was approved by the Research Ethics Office at INRS-EMT and at McGill University (Montreal, Canada).

B. Stimulus

As stimulus, a clean double-sentence utterance commonly used in subjective quality tests was used. In order to generate distorted speech stimuli of varying quality, the clean speech stimulus was convolved with recorded room impulse responses of two different reverberation times (RTs) to generate reverberant speech signals. Table I shows the reverberation times of the room responses and the mean opinion scores (MOS) of the respective speech signals. The content of the speech stimulus was emotionally neutral in order to avoid any influence of content on the perceptual quality. The sentence was uttered by a male speaker in an anechoic chamber and digitized at 8 kHz sampling rate with 16-bit resolution. For consistency, all files were normalized to -26 dBov using the ITU-T P.56 voltmeter [26].

C. Methodology for Subjective Test

Participants were asked to fill a demographic questionnaire, and to perform a subjective quality test. Participants rated the arousal, valence and dominance dimensions using a 9-point continuous SAM pictorial scale. Finally subjects were

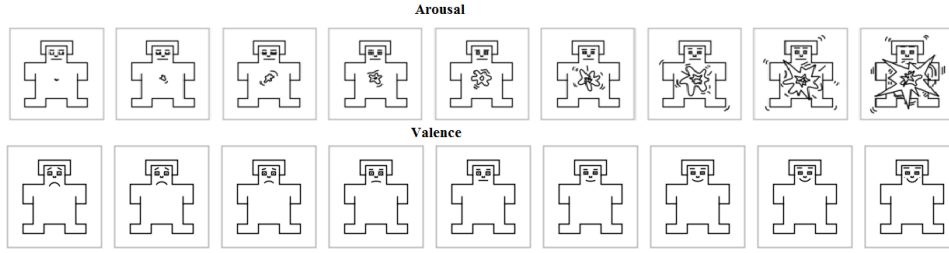


Fig. 2: Self Assessment Manikins (SAM) Scales for Emotion Assessment

TABLE I: Characteristics of Speech Stimuli used in Pilot Experiment

Clean (High Quality)	Medium Quality(MQ)	Low Quality (LQ)
$RT = 0.0s$	$RT = 0.4s$	$RT = 1.5s$
$MOS = 4.2$	$MOS = 3.8$	$MOS = 2.8$

asked to report their experience and emotions with respect to three different quality conditions, namely clean (high quality), medium and low quality. Stimuli were repeated thrice for each condition.

D. Methodology for EEG Experiment

For EEG recording, a 64 channel Biosemi system was used, with electrodes arranged in the 10-20 standard system, as depicted by Figure 3. Four electrodes for electrooculography (EOG) and two for mastoid-electrodes (right and left) were also used as reference. The test consisted of a so-called oddball paradigm; more specifically, the clean speech file served as the so-called “*standard*” stimulus and the reverberant files served as *deviants*. Clean and reverberant speech files were delivered in a pseudo-randomized order, forcing at least one standard stimulus to be presented between successive deviants, in sequences of 100 trials. Stimulus sequences were presented with an inter-stimulus-interval varying from 1000 to 1800 ms. Participants were seated comfortably and were instructed to press a button, whether they detected the clean stimulus or one of the deviants. Stimuli were presented binaurally at the individual’s preferred listening level through in-ear headphones.

1) *EEG Data Pre-processing*: For the EEG analysis, the MATLAB toolbox EEGLAB [27] was used. Data was recorded at 512Hz but down-sampled to 250Hz and band-pass filtered between 1 and 60 Hz for offline analysis. All channels were re-referenced to the average of all EEG-channels. EEG epochs with a length of 3000 ms, time locked to the onset of the stimuli, including a 200 ms pre-stimulus baseline, were extracted and averaged separately for each quality level and for each participant. Movement and eye blink artefacts were removed using visual inspection and independent component analysis (ICA).

2) *EEG Feature*: The EEG $\delta - \beta$ coupling was computed as the Pearson correlation coefficient between the δ and β

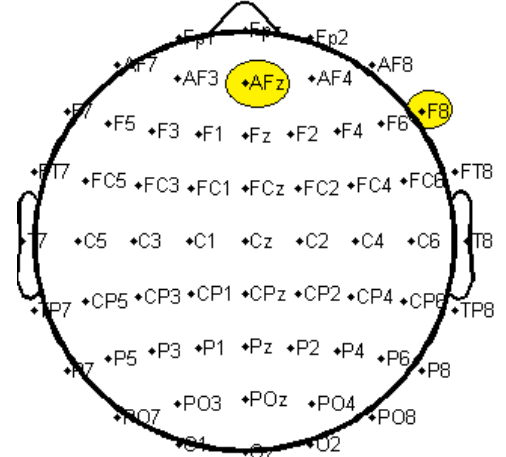


Fig. 3: Electrode Positions based on EEG 10-20 System

frequency band powers for the mid frontal electrode AFz and anterior temporal lobe electrode $F8$ as depicted by Figure 3. Frequency sub-band powers were computed using the Welch’s method. To generate the distribution of coupling coefficients, each subjects’ δ and β powers were correlated across all trials for each quality condition, thus resulting in an overall (averaged over all participants and trials) coupling coefficient for the clean, medium-, and low-quality conditions. Lastly, Fisher’s Z -transformation was used to normalize the coupling coefficients so they could be correlated with the subjective affective ratings of arousal and valence.

IV. EXPERIMENTAL RESULTS

A. Subjective Results

Based on the SAM scale, valence and arousal ratings were calculated across the three different quality conditions. As can be seen from Figure 4, both arousal and valence scores decreased as the speech quality of the stimuli decreased. To better understand the impact of the bi-dimensional (valence-arousal) interaction with respect to three different quality conditions, a valence-arousal graph is shown in Figure 5, where the X-axis represents the SAM scores for valence and the Y-axis for arousal. The data are centered at (5,5), which is the neutral emotional state as per the 9-point SAM scale. The

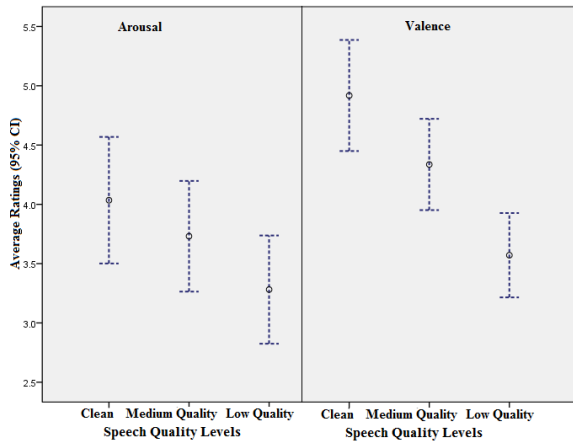


Fig. 4: Subjective Valence and Arousal ratings with Mean and Confidence Intervals

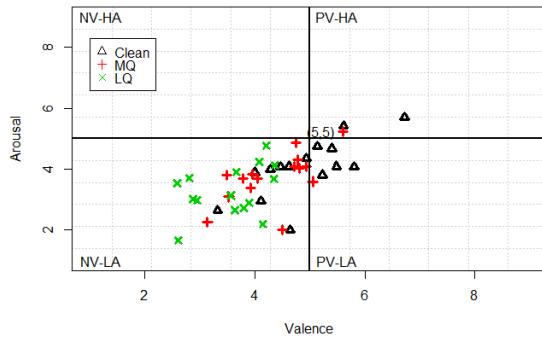


Fig. 5: Arousal-Valence Chart

positive valence and high arousal ($PV - HA$) quadrant represents emotions such as happiness, excitement, and alertness. The positive valence and low arousal ($PV - LA$) quadrant normally represents emotional characteristics like satisfaction, relaxation, and content. The negative valence and high arousal ($NV - HA$) quadrant, in turn, represents emotional characteristics such as agitation, and anger. Affective behaviors such as despair, fatigue, and dissatisfaction are represented in the negative valence and low arousal ($NV - LA$) quadrant. As can be seen, for the clean signal, the majority of the participants rated the stimulus between 4 – 6 in the arousal and valence scales, thus corroborating the neutrality of the speech content. As the speech quality decreases, however, the majority of participants rated between 2 – 4 in the arousal and valence scales, thus pointing towards the negative “affect” quadrants.

To ascertain the significance of the observed differences in valence and arousal with respect to three different quality levels, hypothesis testing with a paired t-test was used. The results are presented in Table II. As can be seen, the average observed difference for arousal and valence ratings were found

TABLE II: Paired T-test of Valence and Arousal scores across three different Quality conditions

Condition Pairs	Valence	Arousal
$Clean - MQ$	$t = 7.2, p \leq 0.05$	$t = 4.2, p \leq 0.05$
$Clean - LQ$	$t = 8.4, p \leq 0.05$	$t = 3.9, p \leq 0.05$
$MQ - LQ$	$t = 8.8, p \leq 0.05$	$t = 3.6, p \leq 0.05$

TABLE III: Pearson Correlation between δ and β powers at EEG electrodes AFz and $F8$ across three different quality conditions

$\delta - \beta$ coupling	Clean	Medium Quality	Low Quality
at AFz	$0.45, p \leq 0.09$	$0.56, p \leq 0.05$	$0.68, p \leq 0.01$
at $F8$	$0.49, p \leq 0.06$	$0.61, p \leq 0.05$	$0.71, p \leq 0.01$

to be statistically different across the three quality conditions, suggesting that emotional states play a significant role in QoE perception.

B. Neurophysiological Results

In line with our expectations, the coupling between $\delta - \beta$ frequency bands at the mid frontal AFz electrode and the anterior right temporal $F8$ electrode increased with decreasing quality conditions, as presented in Table III. This indicates a rise in behavioral inhibition states (possibly such as frustration, dislike, boredom and anxiety) with a decrease in speech quality levels.

When testing the coupling coefficients as correlates of subjective affective ratings at the midfrontal electrode AFz , significant correlations were obtained with arousal as shown in Table IV. A significant Pearson correlation of -0.30 and a significant Spearman rank correlation of -0.34 were obtained. Similarly, a significant negative Pearson correlation of -0.35 and a significant Spearman rank correlation of -0.38 were also obtained between coupling coefficients and subjective valence ratings at the right anterior temporal electrode $F8$.

V. DISCUSSION AND CONCLUSION

This study investigated the hypothesis established in clinical research that an increased EEG $\delta - \beta$ coupling promotes behavioral inhibition states. We also witnessed an increase in $\delta - \beta$ coupling at the midfrontal electrode AFz and the anterior right temporal electrode $F8$ from clean to low quality speech stimulus, thus suggesting that distorted stimuli employed in current study invoked behavioral inhibition characteristics. However, it is possible that either by changing the type of artefact or changing the degradation levels of a stimulus, a different $\delta - \beta$ coupling trend could be achieved.

We also found significant negative correlation between normalized $\delta - \beta$ coupling coefficients and subjective affective (arousal and valence) ratings at the EEG electrode AFz , and $F8$ respectively. The obtained correlation level, however, was not high but it followed the hypothesized inverse relationship trend. This is probably because human subjective judgement

TABLE IV: Correlation between EEG feature and affective ratings [*: $p < 0.05$].

EEG Feature		Arousal	Valence
Normalized Coupling Coefficients		$\rho_{pearson}=-0.30^*$ $\rho_{spearman}=-0.34^*$	$\rho_{pearson}=-0.35^*$ $\rho_{spearman}=-0.38^*$

may not always be exact replica of its objective neurophysiological facts. Additionally, it is also likely due to neutrality of the content utilized for the speech stimuli in the experiment, which may not have evoked strong emotional characteristics. For future experiments, it is suggested that stimuli of varying affective content be used to further investigate the role of the $\delta - \beta$ coupling coefficient on objective QoE perception modeling.

These preliminary insights suggest that in addition to subjective QoE methodologies, complex human neurophysiological processes may be used to characterize human behavioral characteristics. For instance, $\delta - \beta$ coupling provides information about user anxiety, frustration and dissatisfaction. As such, if content and service providers want to ensure rich quality of user experience, the $\delta - \beta$ coupling should not be strong enough to incite negative behavioral characteristics in end users.

ACKNOWLEDGMENT

THF is grateful to the Canadian Foundation for Innovation and the Ministère de l'Éducation, du Loisir et du Sport du Québec for funding the experimental facility and the Ministère du Développement Économique, Innovation et Exportation du Québec and the National Science and Engineering Research Council of Canada for funding the pilot studies. The co-authors from Germany acknowledge funding from the Federal Ministry of Education and Research (Grant FKZ 01GQ0850).

REFERENCES

- [1] K. Laghari, N. Crespi, B. Molina, and C.E. Palau, "Qoe aware service delivery in distributed environment," in *Advanced Information Networking and Applications (WAINA), 2011 IEEE Workshops of International Conference on*, 2011, pp. 837–842.
- [2] K. Laghari and K. Connelly, "Toward total Quality of Experience: A QoE model in a communication ecosystem," *Communications Magazine, IEEE*, vol. 50, no. 4, pp. 58–65, April 2012.
- [3] L. Luigi, P. Vincenzo, and F. Maria, *Advanced Image Processing in Magnetic Resonance Imaging*, Taylor & Francis Group, 2005.
- [4] M. Ferrari, L. Mottola, and V. Quaresima, "Principles, techniques, and limitations of Near Infrared Spectroscopy," *Can. J. Appl. Physiol.*, vol. 29(4), pp. 463–487, 2004.
- [5] K. Laghari, R. Gupta, S. Arndt, J.N. Antons, R. Schleicher, S. Moller, and T.H. Falk, "Neurophysiological experimental facility for Quality of Experience (QoE) assessment," in *QCMAN*, 2013.
- [6] K. Laghari, R. Gupta, S. Arndt, J.N. Antons, R. Schleicher, S. Moller, and T.H. Falk, "Auditory BCIs for visually impaired users: Should developers worry about the quality of text-to-speech readers?," in *International BCI Meeting*, 2013.
- [7] S. Scholler, S. Bosse, M.S. Treder, B. Blankertz, G. Curio, K. Muller, and T. Wiegand, "Toward a direct measure of video quality perception using EEG," *Image Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 2619–2629, 2012.
- [8] S. Arndt, J.N. Antons, R. Schleicher, S. Moller, and G. Curio, "Perception of low-quality videos analyzed by means of Electroencephalography," in *Fourth International Workshop on Quality of Multimedia Experience (QoMEX), AUS-Yarra Valley*, 2012.
- [9] M. Mustafa, S. Guthe, and M. Magnor, "Single trial EEG classification of artifacts in videos," *ACM Trans on Applied Perception*, vol. 9(3), pp. 1201–1215, 2012.
- [10] S. Koelstra, C. Muhl, M. Soleymani, J.S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [11] D.J.L.G. Schutter and J.V. Honk, "Salivary cortisol levels and the coupling of midfrontal delta-beta oscillations," *Int. J. Psychophysiol.*, vol. 55, pp. 127129, 2005.
- [12] ITU-T, "Methods for subjective determination of transmission quality: ITU-T Recommendation P.800," Tech. Rep., International Telecommunication Union, Geneva, Switzerland, 1996.
- [13] ITU-T, "Subjective video quality assessment methods for multimedia applications, Recommendation P.910," Tech. Rep., International Telecommunication Union, Geneva, Switzerland, 2008.
- [14] M.M. Bradley and P.J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J Behav Ther Exp Psychiatry*, vol. 25, pp. 4959, 1994.
- [15] P.J. Lang, "The emotion probe: Studies of motivation and attention," *American Psychologist*, vol. 50(5), pp. 372–385, 1995.
- [16] J.D. Morris, "Observations: SAM, the self assessment manikin, an efficient cross-cultural measurement of emotional response," *Journal of Advertising Research*, Nov 1995.
- [17] K. Jonghwa and E. Andre, "Emotion recognition based on physiological changes in music listening," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.
- [18] S. Kai, J. Yu, Y. Huang, and H. Xiaoqiang, "An improved valence-arousal emotion space for video affective content representation and recognition," in *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, 2009, pp. 566–569.
- [19] J.A. Gray, *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*, Oxford University Press, (1st ed.) edition, 1982.
- [20] D. Hagemann, E. Naumann, A. Lurken, G. Becker, S. Maier, and D. Bartussek, "EEG asymmetry, dispositional mood and personality," *Personality and Individual Differences*, vol. 27, pp. 541–568, 1999.
- [21] M.-M. Mesulam, *Principles of behavioral neurology: Patterns in behavioral neuroanatomy: Association areas, the limbic system, and hemispheric specialization*, Philadelphia: Davis, 1985.
- [22] S. Andres, G. Ende, M. Junghofer, J. Kissler, and D. Wilgruber, Eds., *Understanding Emotions (Progress In Brain Research)*, Elsevier Science, 1st edition, 2006.
- [23] K.L. Phan, S.F. Taylor, R.C. Welsh, L.R. Decker, D.C. Noll, T.E. Nichols, J.C. Britton, and I. Liberzon, "Activation of the medial prefrontal cortex and extended amygdala by individual ratings of emotional arousal: a fMRI study," *Biological Psychiatry*, vol. 53(3), pp. 211–215, 2003.
- [24] J.M.V. Peer, K. Roelofs, and P. Spinoven, "Cortisol administration enhances the coupling of midfrontal delta and beta oscillations," *International Journal of Psychophysiology*, vol. 67, no. 2, pp. 144 – 150, 2008.
- [25] G.G. Knyazev, D.J.L.G. Schutter, and J. V. Honk, "Anxious apprehension increases coupling of delta and beta oscillations," *International Journal of Psychophysiology*, vol. 61, no. 2, pp. 283 – 287, 2006.
- [26] ITU-T, "ITU-T recommendation P.56, objective measurement of active speech level," Tech. Rep., International Telecommunication Union, Geneva, 2011.
- [27] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, 2004.